

Fri Mar 14 10:00:43 2003

Seq2-Peps.res

Page 1

> 0 <
0110 Intelligenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file Seq2-Peps.res made by tport on Fri 28 Feb 103 11:22:23-PST.

Query sequence being compared: US-09-978-385-2 (1-805)
Number of sequences searched: 4
Number of scores above cutoff: 4

Results of the initial comparison of US-09-978-385-2 (1-805) with:
File: US09978385.pep

100-
N -
U -
M -
B -
E -
R -
O -
F -
S -
E -
Q -
U -
N -
E -
C -
S -
SCORE 0 1 89 179 268 358 447 537 626 716 805
STDEV -1

PARAMETERS

Similarity matrix PAM-150 K-tuple 2
Threshold level of sim. 16% Joining penalty 20
Mismatch penalty 5.00 Window size 500
Gap size penalty 0.05
Cutoff score 1
Randomization group 0

SEARCH STATISTICS

Scores: Mean Median Standard Deviation
578 752 383.50
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 2431
Number of sequences searched: 4
Number of scores above cutoff: 4

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name Description Init. Opt. Length Score Score Sig. Frame

1. US-09-978-385-2 Sequence 2, Application US 805 805 805 0.59 0
The list of other best scores is:

Sequence Name Description Length Score Score Init. Opt. Sig. Frame
1. US-09-978-385-2 Sequence 2, Application US 805 752 752 0.45 0
2. US-09-978-385-6 Sequence 2, Application US 805 751 751 0.45 0
3. US-09-978-385-9 Sequence 2, Application US 805 751 751 0.45 0

1. US-09-978-385-2 (1-805) 2. Application US/09978385

Initial Score = 805 Optimized Score = 805 Significance = 0.59
Residue Identity = 100% Matches = 805 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
MSSSWLLSLVAVTAQSTIEEQAKTFLDKFHEADLFYQSLASWNTNTITEENVQNMNAGDKWSAF
MSSSWLLSLVAVTAQSTIEEQAKTFLDKFHEADLFYQSLASWNTNTITEENVQNMNAGDKWSAF
MSSSWLLSLVAVTAQSTIEEQAKTFLDKFHEADLFYQSLASWNTNTITEENVQNMNAGDKWSAF
X 10 20 30 40 50 60 70
LKEOSTLAQMPLOEIOMLVKLOLALQNGSSVSEDSKRLNTILMTSTIYSGKVCNPDPOCELL
LKEOSTLAQMPLOEIOMLVKLOLALQNGSSVSEDSKRLNTILMTSTIYSGKVCNPDPOCELL
LKEOSTLAQMPLOEIOMLVKLOLALQNGSSVSEDSKRLNTILMTSTIYSGKVCNPDPOCELL
X 80 90 100 110 120 130 140
LKEOSTLAQMPLOEIOMLVKLOLALQNGSSVSEDSKRLNTILMTSTIYSGKVCNPDPOCELL
LKEOSTLAQMPLOEIOMLVKLOLALQNGSSVSEDSKRLNTILMTSTIYSGKVCNPDPOCELL
LKEOSTLAQMPLOEIOMLVKLOLALQNGSSVSEDSKRLNTILMTSTIYSGKVCNPDPOCELL

150 160 170 180 190 200 210
EPGLNETMANSIDYNERLWAMESEVSKOLRPLYEYVYKEMARAHNYEDYGRGDEYVNDGYD
EPGLNETMANSIDYNERLWAMESEVSKOLRPLYEYVYKEMARAHNYEDYGRGDEYVNDGYD
EPGLNETMANSIDYNERLWAMESEVSKOLRPLYEYVYKEMARAHNYEDYGRGDEYVNDGYD
EPGLNETMANSIDYNERLWAMESEVSKOLRPLYEYVYKEMARAHNYEDYGRGDEYVNDGYD
X 150 160 170 180 190 200 210
EPGLNETMANSIDYNERLWAMESEVSKOLRPLYEYVYKEMARAHNYEDYGRGDEYVNDGYD
EPGLNETMANSIDYNERLWAMESEVSKOLRPLYEYVYKEMARAHNYEDYGRGDEYVNDGYD
EPGLNETMANSIDYNERLWAMESEVSKOLRPLYEYVYKEMARAHNYEDYGRGDEYVNDGYD

220 230 240 250 260 270 280
YSRQOLIEDVEHTEFEIKPLYEHLAHYVRAKIMNAPYSIPICGLPAHLGDMWGRFTNLISLVPRGOK
YSRQOLIEDVEHTEFEIKPLYEHLAHYVRAKIMNAPYSIPICGLPAHLGDMWGRFTNLISLVPRGOK
YSRQOLIEDVEHTEFEIKPLYEHLAHYVRAKIMNAPYSIPICGLPAHLGDMWGRFTNLISLVPRGOK
YSRQOLIEDVEHTEFEIKPLYEHLAHYVRAKIMNAPYSIPICGLPAHLGDMWGRFTNLISLVPRGOK
X 220 230 240 250 260 270 280
YSRQOLIEDVEHTEFEIKPLYEHLAHYVRAKIMNAPYSIPICGLPAHLGDMWGRFTNLISLVPRGOK
YSRQOLIEDVEHTEFEIKPLYEHLAHYVRAKIMNAPYSIPICGLPAHLGDMWGRFTNLISLVPRGOK
YSRQOLIEDVEHTEFEIKPLYEHLAHYVRAKIMNAPYSIPICGLPAHLGDMWGRFTNLISLVPRGOK

290 300 310 320 330 340 350 360
PRIDVTAMVDQADQRIKREAEKFEFVSGLPNMTQGFWEWSMLTDGNVQKAVCHPTAIDLKGFRLIM
PRIDVTAMVDQADQRIKREAEKFEFVSGLPNMTQGFWEWSMLTDGNVQKAVCHPTAIDLKGFRLIM
PRIDVTAMVDQADQRIKREAEKFEFVSGLPNMTQGFWEWSMLTDGNVQKAVCHPTAIDLKGFRLIM
PRIDVTAMVDQADQRIKREAEKFEFVSGLPNMTQGFWEWSMLTDGNVQKAVCHPTAIDLKGFRLIM
X 290 300 310 320 330 340 350 360
PRIDVTAMVDQADQRIKREAEKFEFVSGLPNMTQGFWEWSMLTDGNVQKAVCHPTAIDLKGFRLIM
PRIDVTAMVDQADQRIKREAEKFEFVSGLPNMTQGFWEWSMLTDGNVQKAVCHPTAIDLKGFRLIM
PRIDVTAMVDQADQRIKREAEKFEFVSGLPNMTQGFWEWSMLTDGNVQKAVCHPTAIDLKGFRLIM

370 380 390 400 410 420 430
CTKVTMDPFLTAHHEMGHIQYDMAVAAQPLLRNGANEHGEHVAEGLMSATPRKHLKSLGSLSPDQEDN
CTKVTMDPFLTAHHEMGHIQYDMAVAAQPLLRNGANEHGEHVAEGLMSATPRKHLKSLGSLSPDQEDN
CTKVTMDPFLTAHHEMGHIQYDMAVAAQPLLRNGANEHGEHVAEGLMSATPRKHLKSLGSLSPDQEDN
CTKVTMDPFLTAHHEMGHIQYDMAVAAQPLLRNGANEHGEHVAEGLMSATPRKHLKSLGSLSPDQEDN
X 370 380 390 400 410 420 430
CTKVTMDPFLTAHHEMGHIQYDMAVAAQPLLRNGANEHGEHVAEGLMSATPRKHLKSLGSLSPDQEDN
CTKVTMDPFLTAHHEMGHIQYDMAVAAQPLLRNGANEHGEHVAEGLMSATPRKHLKSLGSLSPDQEDN
CTKVTMDPFLTAHHEMGHIQYDMAVAAQPLLRNGANEHGEHVAEGLMSATPRKHLKSLGSLSPDQEDN

440 450 460 470 480 490 500
ETEINFLKQALITVGTLPFTYMLEKRWMPFKGEIPDDQMKMKREIIVGVEPVPHDETCDPASLF
ETEINFLKQALITVGTLPFTYMLEKRWMPFKGEIPDDQMKMKREIIVGVEPVPHDETCDPASLF
ETEINFLKQALITVGTLPFTYMLEKRWMPFKGEIPDDQMKMKREIIVGVEPVPHDETCDPASLF
ETEINFLKQALITVGTLPFTYMLEKRWMPFKGEIPDDQMKMKREIIVGVEPVPHDETCDPASLF
X 440 450 460 470 480 490 500
ETEINFLKQALITVGTLPFTYMLEKRWMPFKGEIPDDQMKMKREIIVGVEPVPHDETCDPASLF
ETEINFLKQALITVGTLPFTYMLEKRWMPFKGEIPDDQMKMKREIIVGVEPVPHDETCDPASLF
ETEINFLKQALITVGTLPFTYMLEKRWMPFKGEIPDDQMKMKREIIVGVEPVPHDETCDPASLF

510 520 530 540 550 560 570
HVSNDYSFIRYTRITLYOFQFOFALQAKHGPPLKCDISNTEAGOKLFNMLRLKSEPWILALENVGA
HVSNDYSFIRYTRITLYOFQFOFALQAKHGPPLKCDISNTEAGOKLFNMLRLKSEPWILALENVGA
HVSNDYSFIRYTRITLYOFQFOFALQAKHGPPLKCDISNTEAGOKLFNMLRLKSEPWILALENVGA
HVSNDYSFIRYTRITLYOFQFOFALQAKHGPPLKCDISNTEAGOKLFNMLRLKSEPWILALENVGA
X 510 520 530 540 550 560 570
HVSNDYSFIRYTRITLYOFQFOFALQAKHGPPLKCDISNTEAGOKLFNMLRLKSEPWILALENVGA
HVSNDYSFIRYTRITLYOFQFOFALQAKHGPPLKCDISNTEAGOKLFNMLRLKSEPWILALENVGA
HVSNDYSFIRYTRITLYOFQFOFALQAKHGPPLKCDISNTEAGOKLFNMLRLKSEPWILALENVGA

580 590 600 610 620 630 640
KNMAYRPLNFEPLFTWLDQNKNSFVGSTWSPYAAQSTIKVRSLSKALGAKAYEMNDENYLRSSVA
KNMAYRPLNFEPLFTWLDQNKNSFVGSTWSPYAAQSTIKVRSLSKALGAKAYEMNDENYLRSSVA
KNMAYRPLNFEPLFTWLDQNKNSFVGSTWSPYAAQSTIKVRSLSKALGAKAYEMNDENYLRSSVA
KNMAYRPLNFEPLFTWLDQNKNSFVGSTWSPYAAQSTIKVRSLSKALGAKAYEMNDENYLRSSVA
X 580 590 600 610 620 630 640
KNMAYRPLNFEPLFTWLDQNKNSFVGSTWSPYAAQSTIKVRSLSKALGAKAYEMNDENYLRSSVA
KNMAYRPLNFEPLFTWLDQNKNSFVGSTWSPYAAQSTIKVRSLSKALGAKAYEMNDENYLRSSVA
KNMAYRPLNFEPLFTWLDQNKNSFVGSTWSPYAAQSTIKVRSLSKALGAKAYEMNDENYLRSSVA

```

650      660      670      680      690      700      710      720
YAAKQFLKKNOMLFGEDVAVANLKPRISFNFVYAPKNSDIIIPREVEKAIKMSRSLINDAFRLNDN
|||||
YAAKQFLKKNOMLFGEDVAVANLKPRISFNFVYAPKNSDIIIPREVEKAIKMSRSLINDAFRLNDN
650      660      670      680      690      700      710      720
SLEFLGIPTLPBPNOVPVSIWLVGVYVGVVGIIVTIGTIGIRKKKKKARSGENPYASIDISGENN
|||||
SLEFLGIPTLPBPNOVPVSIWLVGVYVGVVGIIVTIGTIGIRKKKKKARSGENPYASIDISGENN
730      740      750      760      770      780      790
PGFQNTDDVOTSF
|||||
PGFQNTDDVOTSF
800      X
PGFQNTDDVOTSF
|||||
PGFQNTDDVOTSF
800      X

```

2. US-09-978-385-6 Sequence 6, Application US/09978385

Initial Score = 752 Optimized Score = 752 Significance = 0.45
 Residue Identity = 82% Matches = 661 Mismatches = 116
 Gaps = 0 Conservative Substitutions = 28

```

X      10      20      30      40      50      60      70
MSSSSWLLSLVAVTAOSTIEBOAKTFIDKFNEADLFYQSSIASMWNNTITEENYONNNAGDKWSAF
|||||
MSSSSWLLSLVAVTAOSTIEBOAKTFIDKFNEADLFYQSSIASMWNNTITEENYONNNAGDKWSAF
10      20      30      40      50      60      70
X      10      20      30      40      50      60      70
MSSSSWLLSLVAVTAOSTIEBOAKTFIDKFNEADLFYQSSIASMWNNTITEENYONNNAGDKWSAF
|||||
MSSSSWLLSLVAVTAOSTIEBOAKTFIDKFNEADLFYQSSIASMWNNTITEENYONNNAGDKWSAF
80      90      100      110      120      130      140
LKEOSTLQAMPLOEIOMLVYKLOLALOOGSSVLSSEKSKRLNTLNTMTSTYSGKVCNPNROBCILL
|||||
LKEOSTLQAMPLOEIOMLVYKLOLALOOGSSVLSSEKSKRLNTLNTMTSTYSGKVCNPNROBCILL
100      110      120      130      140
YEEOSTKQOSFLOEIOPIIKROLALOOGSSALSADKNKOLNTLNTMTSTYSGKVCNPNROBCILL
|||||
YEEOSTKQOSFLOEIOPIIKROLALOOGSSALSADKNKOLNTLNTMTSTYSGKVCNPNROBCILL
150      160      170      180      190      200      210
EPGGLNEIMANSLDYNERLMAWESMRSEVGKOLRPVYEEVYVLKNEARAHNYEDYGDYWRGDIYVNGVDGYD
|||||
EPGGLNEIMANSLDYNERLMAWESMRSEVGKOLRPVYEEVYVLKNEARAHNYEDYGDYWRGDIYVNGVDGYD
150      160      170      180      190      200      210
EPGGLNEIMANSLDYNERLMAWESMRSEVGKOLRPVYEEVYVLKNEARAHNYEDYGDYWRGDIYVNGVDGYD
|||||
EPGGLNEIMANSLDYNERLMAWESMRSEVGKOLRPVYEEVYVLKNEARAHNYEDYGDYWRGDIYVNGVDGYD
220      230      240      250      260      270      280
YSRGQLEDEVEHFEERIKPLVYHLAAVYRAKLMANVPSYISPIGCLPAHLILGDMGGRFTNLSTLVPRGOK
|||||
YSRGQLEDEVEHFEERIKPLVYHLAAVYRAKLMANVPSYISPIGCLPAHLILGDMGGRFTNLSTLVPRGOK
220      230      240      250      260      270      280
YSRGQLEDEVEHFEERIKPLVYHLAAVYRAKLMANVPSYISPIGCLPAHLILGDMGGRFTNLSTLVPRGOK
|||||
YSRGQLEDEVEHFEERIKPLVYHLAAVYRAKLMANVPSYISPIGCLPAHLILGDMGGRFTNLSTLVPRGOK
290      300      310      320      330      340      350      360
PNIQVTDAMAVDADAQRIFFKAERFVSVGLPRMTOGFENSMILDPGVNOKAVCHPTAMDLGKGDPRILM
|||||
PNIQVTDAMAVDADAQRIFFKAERFVSVGLPRMTOGFENSMILDPGVNOKAVCHPTAMDLGKGDPRILM
290      300      310      320      330      340      350      360
PNIQVTDAMAVDADAQRIFFKAERFVSVGLPRMTOGFENSMILDPGVNOKAVCHPTAMDLGKGDPRILM
|||||
PNIQVTDAMAVDADAQRIFFKAERFVSVGLPRMTOGFENSMILDPGVNOKAVCHPTAMDLGKGDPRILM
370      380      390      400      410      420      430
CTKYVMDPFLAAHHEMGHIQDYAAVAQAPFLNRNGANEGFHEAVGEIMLSAATPKHLKSLGILSPDOEDN
|||||
CTKYVMDPFLAAHHEMGHIQDYAAVAQAPFLNRNGANEGFHEAVGEIMLSAATPKHLKSLGILSPDOEDN
370      380      390      400      410      420      430
CTKYVMDPFLAAHHEMGHIQDYAAVAQAPFLNRNGANEGFHEAVGEIMLSAATPKHLKSLGILSPDOEDN
|||||
CTKYVMDPFLAAHHEMGHIQDYAAVAQAPFLNRNGANEGFHEAVGEIMLSAATPKHLKSLGILSPDOEDN
440      450      460      470      480      490      500
ETEINFLKQALITVGLTPPTVYMLEKRWNVFRGELIPKQOMKRWMEKKREIVGVVPRVPHDETYCDPASLF
|||||
ETEINFLKQALITVGLTPPTVYMLEKRWNVFRGELIPKQOMKRWMEKKREIVGVVPRVPHDETYCDPASLF
440      450      460      470      480      490      500
ETEINFLKQALITVGLTPPTVYMLEKRWNVFRGELIPKQOMKRWMEKKREIVGVVPRVPHDETYCDPASLF
|||||
ETEINFLKQALITVGLTPPTVYMLEKRWNVFRGELIPKQOMKRWMEKKREIVGVVPRVPHDETYCDPASLF
510      520      530      540      550      560      570
HVSNDVSFIRYTRTYLQFOFOEALCOAAKHEGPIHAKGDISNSTEAOQKLENNMLRIKSEKSPPTALALENVGA
|||||
HVSNDVSFIRYTRTYLQFOFOEALCOAAKHEGPIHAKGDISNSTEAOQKLENNMLRIKSEKSPPTALALENVGA
510      520      530      540      550      560      570
HVSNDVSFIRYTRTYLQFOFOEALCOAAKHEGPIHAKGDISNSTEAOQKLENNMLRIKSEKSPPTALALENVGA
|||||
HVSNDVSFIRYTRTYLQFOFOEALCOAAKHEGPIHAKGDISNSTEAOQKLENNMLRIKSEKSPPTALALENVGA

```

3. US-09-978-385-2 (1-805) Sequence 9, Application US/09978385

Initial Score = 751 Optimized Score = 751 Significance = 0.45
 Residue Identity = 81% Matches = 659 Mismatches = 118
 Gaps = 0 Conservative Substitutions = 28

```

X      10      20      30      40      50      60      70
MSSSSWLLSLVAVTAOSTIEBOAKTFIDKFNEADLFYQSSIASMWNNTITEENYONNNAGDKWSAF
|||||
MSSSSWLLSLVAVTAOSTIEBOAKTFIDKFNEADLFYQSSIASMWNNTITEENYONNNAGDKWSAF
10      20      30      40      50      60      70
X      10      20      30      40      50      60      70
MSSSSWLLSLVAVTAOSTIEBOAKTFIDKFNEADLFYQSSIASMWNNTITEENYONNNAGDKWSAF
|||||
MSSSSWLLSLVAVTAOSTIEBOAKTFIDKFNEADLFYQSSIASMWNNTITEENYONNNAGDKWSAF
80      90      100      110      120      130      140
LKEOSTLQAMPLOEIOMLVYKLOLALOOGSSVLSSEKSKRLNTLNTMTSTYSGKVCNPNROBCILL
|||||
LKEOSTLQAMPLOEIOMLVYKLOLALOOGSSVLSSEKSKRLNTLNTMTSTYSGKVCNPNROBCILL
100      110      120      130      140
YEEOSTKQOSFLOEIOPIIKROLALOOGSSALSADKNKOLNTLNTMTSTYSGKVCNPNROBCILL
|||||
YEEOSTKQOSFLOEIOPIIKROLALOOGSSALSADKNKOLNTLNTMTSTYSGKVCNPNROBCILL
150      160      170      180      190      200      210
EPGGLNEIMANSLDYNERLMAWESMRSEVGKOLRPVYEEVYVLKNEARAHNYEDYGDYWRGDIYVNGVDGYD
|||||
EPGGLNEIMANSLDYNERLMAWESMRSEVGKOLRPVYEEVYVLKNEARAHNYEDYGDYWRGDIYVNGVDGYD
150      160      170      180      190      200      210
EPGGLNEIMANSLDYNERLMAWESMRSEVGKOLRPVYEEVYVLKNEARAHNYEDYGDYWRGDIYVNGVDGYD
|||||
EPGGLNEIMANSLDYNERLMAWESMRSEVGKOLRPVYEEVYVLKNEARAHNYEDYGDYWRGDIYVNGVDGYD
220      230      240      250      260      270      280
YSRGQLEDEVEHFEERIKPLVYHLAAVYRAKLMANVPSYISPIGCLPAHLILGDMGGRFTNLSTLVPRGOK
|||||
YSRGQLEDEVEHFEERIKPLVYHLAAVYRAKLMANVPSYISPIGCLPAHLILGDMGGRFTNLSTLVPRGOK
220      230      240      250      260      270      280
YSRGQLEDEVEHFEERIKPLVYHLAAVYRAKLMANVPSYISPIGCLPAHLILGDMGGRFTNLSTLVPRGOK
|||||
YSRGQLEDEVEHFEERIKPLVYHLAAVYRAKLMANVPSYISPIGCLPAHLILGDMGGRFTNLSTLVPRGOK
290      300      310      320      330      340      350      360
PNIQVTDAMAVDADAQRIFFKAERFVSVGLPRMTOGFENSMILDPGVNOKAVCHPTAMDLGKGDPRILM
|||||
PNIQVTDAMAVDADAQRIFFKAERFVSVGLPRMTOGFENSMILDPGVNOKAVCHPTAMDLGKGDPRILM
290      300      310      320      330      340      350      360
PNIQVTDAMAVDADAQRIFFKAERFVSVGLPRMTOGFENSMILDPGVNOKAVCHPTAMDLGKGDPRILM
|||||
PNIQVTDAMAVDADAQRIFFKAERFVSVGLPRMTOGFENSMILDPGVNOKAVCHPTAMDLGKGDPRILM
370      380      390      400      410      420      430
CTKYVMDPFLAAHHEMGHIQDYAAVAQAPFLNRNGANEGFHEAVGEIMLSAATPKHLKSLGILSPDOEDN
|||||
CTKYVMDPFLAAHHEMGHIQDYAAVAQAPFLNRNGANEGFHEAVGEIMLSAATPKHLKSLGILSPDOEDN
370      380      390      400      410      420      430
CTKYVMDPFLAAHHEMGHIQDYAAVAQAPFLNRNGANEGFHEAVGEIMLSAATPKHLKSLGILSPDOEDN
|||||
CTKYVMDPFLAAHHEMGHIQDYAAVAQAPFLNRNGANEGFHEAVGEIMLSAATPKHLKSLGILSPDOEDN
440      450      460      470      480      490      500
ETEINFLKQALITVGLTPPTVYMLEKRWNVFRGELIPKQOMKRWMEKKREIVGVVPRVPHDETYCDPASLF
|||||
ETEINFLKQALITVGLTPPTVYMLEKRWNVFRGELIPKQOMKRWMEKKREIVGVVPRVPHDETYCDPASLF
440      450      460      470      480      490      500
ETEINFLKQALITVGLTPPTVYMLEKRWNVFRGELIPKQOMKRWMEKKREIVGVVPRVPHDETYCDPASLF
|||||
ETEINFLKQALITVGLTPPTVYMLEKRWNVFRGELIPKQOMKRWMEKKREIVGVVPRVPHDETYCDPASLF
510      520      530      540      550      560      570
HVSNDVSFIRYTRTYLQFOFOEALCOAAKHEGPIHAKGDISNSTEAOQKLENNMLRIKSEKSPPTALALENVGA
|||||
HVSNDVSFIRYTRTYLQFOFOEALCOAAKHEGPIHAKGDISNSTEAOQKLENNMLRIKSEKSPPTALALENVGA
510      520      530      540      550      560      570
HVSNDVSFIRYTRTYLQFOFOEALCOAAKHEGPIHAKGDISNSTEAOQKLENNMLRIKSEKSPPTALALENVGA
|||||
HVSNDVSFIRYTRTYLQFOFOEALCOAAKHEGPIHAKGDISNSTEAOQKLENNMLRIKSEKSPPTALALENVGA

```



```

Db      541 KCDISNSTEAGOKLFENMLRLGKSEPTWLALENVYGAKNNMVRPLNYPFELFTWLKDQNK 600
        |||
QY      601 NSDVGSTMSGYAAOOSIKVRSLSKLSALGDKAAYMDNEWAYLRRSSVAAMAQYFLKYKN 660
        |||
Db      601 NSSVGSTMSGYAAOOSIKVRSLSKLSALGDKAAYMDNEWAYLRRSSVAAMAQYFLKYKN 660
        |||
QY      661 OMILFEEDRYRANLKPRISFNFFVTAPKNVSDIIPTEVEKAIIRMSRINDAFRLND 720
        |||
Db      661 OMILFEEDRYRANLKPRISFNFFVTAPKNVSDIIPTEVEKAIIRMSRINDAFRLND 720
        |||
QY      721 SLEFGIOPTLGPPNPQPVYSIMLVFGVNGVIYWGIVILIIFTGIDRRKKRKARGENP 780
        |||
Db      721 SLEFGIOPTLGPPNPQPVYSIMLVFGVNGVIYWGIVILIIFTGIDRRKKRKARGENP 780
        |||
QY      781 YASIDISKGENNPGQNITDDVQTSTF 805
        |||
Db      781 YASIDISKGENNPGQNITDDVQTSTF 805
        |||

RESULT 2
AA67310
ID      AA67310 standard; Protein; 805 AA.
AC      AAY67310;
XX
XX
DT      11-APR-2000 (first entry)
XX
DE      Human MPROT15 amino acid sequence #1.
XX
XX      MPROT15; treatment; hypertension; human; myocardial disease; apoplexy;
KW      heart disease; apoplexy; heart disease; nervous denaturation; hormone;
KW      Alzheimer's disease; cytokine.
XX
XX      Homo sapiens.
OS
PN      JP11318472-A.
XX
PD      24-NOV-1999.
XX
PF      22-JAN-1999; 99JP-0014949.
XX
PR      13-MAY-1998; 98GB-0010373.
PR      18-AUG-1998; 98GB-0018009.
PA      (SMIK ) SMITHKLIN BEECHAM PLC.
XX
XX      WPI; 2000-109268/10.
DR      N-PSDB; AAZ59465.
XX
XX      MPROT15 polypeptide and MPROT15 polynucleotides - useful for the
PT      treatment of hypertension, myocardial diseases, apoplexy, heart
PT      diseases, nervous denaturation, Alzheimer's disease etc.
XX
XX      Claim 1; Page 15; 22pp; Japanese.
PS
XX
CC      This is amino acid sequence #1 of human MPROT15. The MPROT15
CC      polynucleotide and polypeptide sequences can be used for the treatment of
CC      hypertension, myocardial diseases, apoplexy, heart diseases, nervous
CC      denaturation, Alzheimer's disease and diseases related to the processing
CC      of peptide hormones and cytokines.
XX
XX      Sequence 805 AA:
SO

Query Match          100.0%; Score 4291; DB 21; Length 805;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MSSSSWLLSLAVNTAOSTIEBQAETLDKFNFHAEDELFYQSLSAMNNTNITEENVQ 60
        |||
Db      1 MSSSSWLLSLAVNTAOSTIEBQAETLDKFNFHAEDELFYQSLSAMNNTNITEENVQ 60
        |||

61 NMNNAAGKMSAFLEKEOSTIAQMPYLAQEIONLTVKIQLOALQQNGSSVYLSEDKSKRLNTLI 120

```


QY 361 CTCTVMDPFLTAHHEMGIQYDMAYAAOPELLRNGANGFHEAVGEIMSLSAATPKHLKS 420
 DB 361 CTCTVMDPFLTAHHEMGIQYDMAYAAOPELLRNGANGFHEAVGEIMSLSAATPKHLKS 420
 QY 421 IGLSPDFOEDNETEINFLKQALITVGTLPPTMLEKRWMTVEKGEIPKDOMKKWEM 480
 DB 421 IGLSPDFOEDNETEINFLKQALITVGTLPPTMLEKRWMTVEKGEIPKDOMKKWEM 480
 QY 481 KREIVGVPEVPHDETCYDPAFLFHSNDYSFIRYRTYTLVQFOFQALCOAAKHEBPLH 540
 DB 481 KREIVGVPEVPHDETCYDPAFLFHSNDYSFIRYRTYTLVQFOFQALCOAAKHEBPLH 540
 QY 541 KCDISNSTEAGOKLFNMLRLKSEPWTLALENVGAKMMNVRPLNFEPLFTWLKQNK 600
 DB 541 KCDISNSTEAGOKLFNMLRLKSEPWTLALENVGAKMMNVRPLNFEPLFTWLKQNK 600
 QY 601 NSFVGSTDMSPYADQSIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVK 660
 DB 601 NSFVGSTDMSPYADQSIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVK 660
 QY 661 QMILFGEEDYRVANLKRISFNFFVTAPKNVSDIIPREVEKAIKMSRSRINDAPFLNDN 720
 DB 661 QMILFGEEDYRVANLKRISFNFFVTAPKNVSDIIPREVEKAIKMSRSRINDAPFLNDN 720
 QY 721 SLEFLGIQPTLGPNOFPVSIWLVGVYGVVGVIVLITGTIRDKKKKKASGENP 780
 DB 721 SLEFLGIQPTLGPNOFPVSIWLVGVYGVVGVIVLITGTIRDKKKKKASGENP 780
 QY 781 YASIDISKGNNGFQONTDVCYTSF 805
 DB 781 YASIDISKGNNGFQONTDVCYTSF 805

RESULT 4
 AAB48095
 ID AAB48095 standard; Protein: 805 AA.
 AC AAB48095;

DT 19-MAR-2001 (first entry)

DE Human Zace2 protein.

XX Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
 KW zinc metalloproteinase; blood pressure; zinc protease; hypertension;
 KW ventricular systolic dysfunction; renal impairment; heart failure;
 KW scleroderma renal crisis; atherosclerosis; antinflammatory; human;
 KW antithrombotic; bradykinin inactivator.

XX Homo sapiens.

XX MO200070032-A1.

XX 23-NOV-2000.

XX 03-MAY-2000; 2000MO-0511932.

XX 13-MAY-1999; 99US-0311482.

XX 27-AUG-1999; 99US-0384706.

XX (ZIMO) ZYMOGENETICS INC.

PI Pladdington CS, Petrie CR, Shoemaker KE, Bishop PD;

DR WPI; 2001-025018/03.

DR N-PSDB; AAC84366, AAC84367.

PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 XX associated with inflammation such as arthritis and enterocolitis -
 PS Example 1; Page 95-100; 125pp; English.

CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood
 CC pressure regulation and fertility. Zace2 can be expressed by standard
 CC recombinant methodology. Zace2 polypeptides are useful for treating an
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
 CC diseases associated with inflammation like arthritis and enterocolitis,
 CC as targets for identifying modulators of zinc protease activity, for
 CC screening or identifying new angiotensin-converting enzyme (ACE)
 CC inhibitors, and as a basis for rational drug design for inhibitory
 CC molecules. The nucleic acids can be used to detect the expression of a
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and
 CC for detecting and localizing Zace2 gene expression in tissue samples,
 CC to determine whether a subject's chromosomes contain a mutation in the
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.
 CC Inhibitors of ACE are used for treating hypertension of various
 CC conditions, including left ventricular systolic dysfunction, progressive
 CC renal impairment, scleroderma renal crisis, congestive heart failure due
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
 CC used to treat infertility while Zace2 antagonists are used for inducing
 CC infertility. The present sequence represents the human Zace2 protein.
 XX

Sequence 805 AA;

Query Match 100.0%; Score 4291; DB 22; Length 805;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSWLLSIVAVTAOSTIEQAKTFPLDKRHHAEPLFOSSLASNNVTNTEENVQ 60
 DB 1 MSSSSWLLSIVAVTAOSTIEQAKTFPLDKRHHAEPLFOSSLASNNVTNTEENVQ 60
 QY 61 MNMAGDKWSAFLEKQSTLAQMYPLQETQNLVQLQALQOQSSVYSEKSKRLTIL 120
 DB 61 MNMAGDKWSAFLEKQSTLAQMYPLQETQNLVQLQALQOQSSVYSEKSKRLTIL 120
 QY 121 NTMSTIYSTGKVCNPDNPOECLEPGLNETMANSLDYNERLMAESWSEVQKQRLPY 180
 DB 121 NTMSTIYSTGKVCNPDNPOECLEPGLNETMANSLDYNERLMAESWSEVQKQRLPY 180
 QY 181 EBYVVLKNDMARANHYEDYGDYRGDYEVDYDSRQGLLEDYHFEETKRLPYELH 240
 DB 181 EBYVVLKNDMARANHYEDYGDYRGDYEVDYDSRQGLLEDYHFEETKRLPYELH 240
 QY 241 HAYYRAKIMNAYPYXISIPICLPAHLGDMWGRWNLXSLYVFEQKKNIDVTDAVDQ 300
 DB 241 HAYYRAKIMNAYPYXISIPICLPAHLGDMWGRWNLXSLYVFEQKKNIDVTDAVDQ 300
 QY 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWENSMLTDEGNVQKAVCHPTAMDLCGDFRILM 360
 DB 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWENSMLTDEGNVQKAVCHPTAMDLCGDFRILM 360
 QY 361 CTCTVMDPFLTAHHEMGIQYDMAYAAOPELLRNGANGFHEAVGEIMSLSAATPKHLKS 420
 DB 361 CTCTVMDPFLTAHHEMGIQYDMAYAAOPELLRNGANGFHEAVGEIMSLSAATPKHLKS 420
 QY 421 IGLSPDFOEDNETEINFLKQALITVGTLPPTMLEKRWMTVEKGEIPKDOMKKWEM 480
 DB 421 IGLSPDFOEDNETEINFLKQALITVGTLPPTMLEKRWMTVEKGEIPKDOMKKWEM 480
 QY 481 KREIVGVPEVPHDETCYDPAFLFHSNDYSFIRYRTYTLVQFOFQALCOAAKHEBPLH 540
 DB 481 KREIVGVPEVPHDETCYDPAFLFHSNDYSFIRYRTYTLVQFOFQALCOAAKHEBPLH 540
 QY 541 KCDISNSTEAGOKLFNMLRLKSEPWTLALENVGAKMMNVRPLNFEPLFTWLKQNK 600
 DB 541 KCDISNSTEAGOKLFNMLRLKSEPWTLALENVGAKMMNVRPLNFEPLFTWLKQNK 600
 QY 601 NSFVGSTDMSPYADQSIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVK 660
 DB 601 NSFVGSTDMSPYADQSIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVK 660
 QY 661 QMILFGEEDYRVANLKRISFNFFVTAPKNVSDIIPREVEKAIKMSRSRINDAPFLNDN 720
 DB 661 QMILFGEEDYRVANLKRISFNFFVTAPKNVSDIIPREVEKAIKMSRSRINDAPFLNDN 720

Db 661 QMILFGEEDVAVANLKRISFNFPVAPKAVSDIIPTEVEKAIIRMSRSRINDAFRLNDN 720

Oy 721 SLEFLGIQPLTGPNNOPPVSIWLVGVGVIVGIVILITGIRDRKKKKRARGENP 780

Db 721 SLEFLGIQPLTGPNNOPPVSIWLVGVGVIVGIVILITGIRDRKKKKRARGENP 780

Oy 781 YASIDISKGENNGPONTDVOVSF 805

Db 781 YASIDISKGENNGPONTDVOVSF 805

RESULT 5

AAU99701 ID AAU99701 standard; protein; 805 AA.

XX AAU99701;

XX 24-SEP-2002 (first entry)

DE Human angiotensin converting enzyme-2 (ACE-2) protein.

XX Human; angiotensin converting enzyme-2; ACE-2; body weight disorder;

KW muscle mass; body fat; obesity; diabetes; atherosclerosis; weight loss;

KW lipid metabolism; weight gain; anorexia; cachexia; bulimia; sepsis;

KW familial partial lipodystrophy; hypercholesterolemia; hyperlipidaemia;

KW aberrant metabolic rate; heart failure; left ventricular hypertrophy;

KW neurodegenerative disorder; peptide hormone; cytokine processing;

KW myocardial infarction; cardiomyopathy; inflammatory bowel disease;

KW systemic inflammation response syndrome; polytrauma; pain; stroke;

KW bone destruction; rheumatoid arthritis; osteoarthritis; asthma;

KW periodontal disease; dysmenorrhoea; premature labour; brain oedema;

KW focal injury; diffuse axonal injury; reperfusion injury; scar formation;

KW cerebral vasospasm; subarachnoid haemorrhage; allergic disorder;

KW adult respiratory distress syndrome; wound healing; appetite;

KW body mass index.

OS Homo sapiens.

XX Key: Location/Qualifiers

FT Peptide 1..18

FT /label= Signal_peptide

FT 19..805

FT Protein /label= Mature_human_ACE_2_protein

FT WO200239997-A2.

XX 23-MAY-2002.

PD 31-OCT-2001; 2001WO-US45703.

PF 01-NOV-2000; 2000US-0794216

PR 29-MAY-2001; 2001US-0870382

PR 19-OCT-2001; 2001US-371741P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Action SL, Ocain TD, Gould AE, Dales NA, Guan B, Brown JA;

PI Palane M, Kadambi VJ, Solomon M, Stricker-Krongrad A;

XX MPI: 2002-547572/58.

DR N-PSDB; ABR67623.

XX Treating body weight disorder and increasing muscle mass comprises

PT administering angiotensin converting enzyme-2 modulating compound

XX Example 5; Page 387-390; 395pp; English.

XX The present invention describes a new method of treating a body weight

CC disorder, increasing muscle mass and decreasing body fat by

CC administration of angiotensin converting enzyme (ACE)-2 modulating

CC compound. The invention can be used for treating body weight disorders,

CC particularly obesity of at least grade 1, diabetes, atherosclerosis and

CC a state associated with lipid metabolism. The method is used for treating

CC rapid weight loss, rapid weight gain, anorexia, cachexia, bulimia,

CC generalised partial lipodystrophy, familial partial lipodystrophy,

CC hypercholesterolemia, hyperlipidaemia, an aberrant metabolic rate,

CC congestive heart failure, chronic heart failure, left ventricular

CC hypertrophy, acute heart failure, neurodegenerative disorders (e.g.

CC Alzheimer's disease, Parkinson's disease and Huntington's disease),

CC diseases associated with peptide hormones or cytokine processing,

CC myocardial infarction, cardiomyopathy, systemic inflammation response

CC syndrome, sepsis, polytrauma, inflammatory bowel disease, acute and

CC chronic pain, bone destruction in rheumatoid arthritis and osteoarthritis

CC and periodontal disease, dysmenorrhoea, premature labour, brain oedema

CC following focal injury, diffuse axonal injury, stroke, reperfusion

CC injury, cerebral vasospasm after subarachnoid haemorrhage, allergic

CC disorders including asthma, adult respiratory distress syndrome, wound

CC healing and scar formation. The invention decreases the appetite,

CC increases muscle mass and decreases body fat of subject having body mass

CC index of greater than 23 (preferably 24.9)kg/m². The present amino

CC acid sequence represents the human ACE-2 protein of the invention.

XX

Sequence 805 AA:

Query Match 100.0%; Score 4291; DB 23; Length 805;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MSSSSMLLSLVAVTAOSTIEEOAKTFLDKFHEADELFYOSLSAMYNNTITEENVQ 60

Db 1 MSSSSMLLSLVAVTAOSTIEEOAKTFLDKFHEADELFYOSLSAMYNNTITEENVQ 60

Oy 61 MNMAGDKMSAFLEKOSTLAQMYPLQEIQNTLVKIQLOALQONGSSVSEDKSKRLNTLL 120

Db 61 MNMAGDKMSAFLEKOSTLAQMYPLQEIQNTLVKIQLOALQONGSSVSEDKSKRLNTLL 120

Oy 121 NTMSTIYSTKVCNPNPPOCCLLEFGINFTANSIDYNERLAWMSSEVKOLRPLY 180

Db 121 NTMSTIYSTKVCNPNPPOCCLLEFGINFTANSIDYNERLAWMSSEVKOLRPLY 180

Oy 121 NTMSTIYSTKVCNPNPPOCCLLEFGINFTANSIDYNERLAWMSSEVKOLRPLY 180

Db 121 NTMSTIYSTKVCNPNPPOCCLLEFGINFTANSIDYNERLAWMSSEVKOLRPLY 180

Oy 181 EEEYVLKKNEMARANHEDYDYGWRCDEYGVGYDSRGCLIEDVHTEETKPLYEHL 240

Db 181 EEEYVLKKNEMARANHEDYDYGWRCDEYGVGYDSRGCLIEDVHTEETKPLYEHL 240

Oy 181 EEEYVLKKNEMARANHEDYDYGWRCDEYGVGYDSRGCLIEDVHTEETKPLYEHL 240

Db 181 EEEYVLKKNEMARANHEDYDYGWRCDEYGVGYDSRGCLIEDVHTEETKPLYEHL 240

Oy 241 HAYVRAKIMNAYPSYISPIGCLPAHLIGDMWGRFTNLYSLYVPGCKRPIDVTAMVQ 300

Db 241 HAYVRAKIMNAYPSYISPIGCLPAHLIGDMWGRFTNLYSLYVPGCKRPIDVTAMVQ 300

Oy 301 AMDAQRIFKFAKRFVSGLPNMTQGFWENSMILDPENYOKAVCHPTAMDLGKGRRLIM 360

Db 301 AMDAQRIFKFAKRFVSGLPNMTQGFWENSMILDPENYOKAVCHPTAMDLGKGRRLIM 360

Oy 301 AMDAQRIFKFAKRFVSGLPNMTQGFWENSMILDPENYOKAVCHPTAMDLGKGRRLIM 360

Db 301 AMDAQRIFKFAKRFVSGLPNMTQGFWENSMILDPENYOKAVCHPTAMDLGKGRRLIM 360

Oy 361 CTKYTMDDFLTAHHEMGIQDYMAVAAQPLRLRNGANEGFHEAVGEIMLSAATPKHLNS 420

Db 361 CTKYTMDDFLTAHHEMGIQDYMAVAAQPLRLRNGANEGFHEAVGEIMLSAATPKHLNS 420

Oy 361 CTKYTMDDFLTAHHEMGIQDYMAVAAQPLRLRNGANEGFHEAVGEIMLSAATPKHLNS 420

Db 361 CTKYTMDDFLTAHHEMGIQDYMAVAAQPLRLRNGANEGFHEAVGEIMLSAATPKHLNS 420

Oy 421 IGLSPDFQEDNTEINFLKQALTVGLTPTYMLKRWMYFKGIPIDOMKKWEM 480

Db 421 IGLSPDFQEDNTEINFLKQALTVGLTPTYMLKRWMYFKGIPIDOMKKWEM 480

Oy 421 IGLSPDFQEDNTEINFLKQALTVGLTPTYMLKRWMYFKGIPIDOMKKWEM 480

Db 421 IGLSPDFQEDNTEINFLKQALTVGLTPTYMLKRWMYFKGIPIDOMKKWEM 480

Oy 481 KREIVGVVEPVPHEDETCDPASLPHVSNDSFIRYRTLLYOPFOALCOAKHNEGPLH 540

Db 481 KREIVGVVEPVPHEDETCDPASLPHVSNDSFIRYRTLLYOPFOALCOAKHNEGPLH 540

Oy 481 KREIVGVVEPVPHEDETCDPASLPHVSNDSFIRYRTLLYOPFOALCOAKHNEGPLH 540

Db 481 KREIVGVVEPVPHEDETCDPASLPHVSNDSFIRYRTLLYOPFOALCOAKHNEGPLH 540

Oy 541 KCDISNSTRGAKLFNMLRLKSEPTLLALENVGAKNMVRLPLNFEPLFTWLKQDNK 600

Db 541 KCDISNSTRGAKLFNMLRLKSEPTLLALENVGAKNMVRLPLNFEPLFTWLKQDNK 600

Oy 541 KCDISNSTRGAKLFNMLRLKSEPTLLALENVGAKNMVRLPLNFEPLFTWLKQDNK 600

Db 541 KCDISNSTRGAKLFNMLRLKSEPTLLALENVGAKNMVRLPLNFEPLFTWLKQDNK 600

Oy 601 NSFVGWSTDMSPYADQSIKVRISLKSALGDKAYEMNENEMTLFRSSVAYAMROFLFKVN 660

Db 601 NSFVGWSTDMSPYADQSIKVRISLKSALGDKAYEMNENEMTLFRSSVAYAMROFLFKVN 660

Oy 601 NSFVGWSTDMSPYADQSIKVRISLKSALGDKAYEMNENEMTLFRSSVAYAMROFLFKVN 660

Db 601 NSFVGWSTDMSPYADQSIKVRISLKSALGDKAYEMNENEMTLFRSSVAYAMROFLFKVN 660

Oy 661 QMILFGEEDVAVANLKRISFNFPVAPKAVSDIIPTEVEKAIIRMSRSRINDAFRLNDN 720

Db 661 QMILFGEEDVAVANLKRISFNFPVAPKAVSDIIPTEVEKAIIRMSRSRINDAFRLNDN 720

QY 721 SLEFLGIQPTLGPNNPVSIMLIVGVGVVIGVILITFTGIRDRKKKKKASGENP 780
 DB 721 SLEFLGIQPTLGPNNPVSIMLIVGVGVVIGVILITFTGIRDRKKKKKASGENP 780
 QY 781 YASIDISKGNNPGFQNTDDVQTSF 805
 DB 781 YASIDISKGNNPGFQNTDDVQTSF 805

RESULT 6
 AAE20353
 ID AAE20353 standard; Protein; 805 AA;
 AC AAE20353;
 XX
 XX 18-JUN-2002 (first entry)
 DE Human ACE-2 full-length protein.
 KW Human; angiotensin converting enzyme-2; ACE-2; therapy; hypertension;
 KW peptidyl dipeptidase A; blood pressure; hypotension; atherosclerosis;
 KW myocardial infarction; heart failure; arrhythmia; renal failure; gene;
 KW inflammation; fertility; enzyme; EC 3.4.15.1.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT /label= Signal_peptide
 FT Protein
 FT /note= "Mature ACE-2 protein"
 FT Domain
 FT /note= "Extracellular domain"
 FT /note= "Zinc binding domain (ZBD)"
 FT Domain
 FT /note= "Transmembrane domain"
 FT /note= "766..805"
 FT Domain
 FT /note= "Cytoplasmic domain"
 XX
 PN WO200212471-A2.
 PD 14-FEB-2002.
 XX
 PD 09-AUG-2001; 2001WO-US25059.
 XX
 PR 09-AUG-2000; 2000US-0635501.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Acton S, Robison KE, Hsieh FY;
 XX
 DR WPI; 2002-257481/30.
 DR N-PSDB; AAD32586.
 XX
 PT Isolated human polypeptide, known as angiotensin converting enzyme-2,
 PT useful for treating or preventing the development of an abnormal blood
 PT pressure or related diseases, e.g. hypertension, heart failure or
 PT myocardial infarction -
 XX
 XX Claim 2; Fig 1; 218P; English.
 PS
 PS
 CC The invention relates to human angiotensin converting enzyme-2 (ACE-2)
 CC polypeptides and polynucleotides. ACE-2 is also known as peptidyl
 CC dipeptidase A (EC 3.4.15.1). Polypeptides of the invention are useful
 CC for treating or preventing the development of abnormal blood pressure
 CC and diseases or disorders associated with the protein in a subject. The
 CC diseases include hypertension, hypotension, congestive heart failure,
 CC chronic heart failure, acute heart failure, myocardial infarction,
 CC atherosclerosis, arrhythmia and renal failure. They are also useful
 CC for treating inflammatory conditions and diseases relating to fertility.
 CC The present sequence is human full-length ACE-2 protein.
 XX

SQ Sequence 805 AA;
 Query Match 100.0%; Score 4291; DB 23; Length 805;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSLLSLVAVTAOSTIEBOAKTFLDKFNHAEEDLFYQSSLASMNVNTNTEENVQ 60
 DB 1 MSSSSLLSLVAVTAOSTIEBOAKTFLDKFNHAEEDLFYQSSLASMNVNTNTEENVQ 60
 QY 61 NNNNAGDKWSAFLEKQSTLAQMPLOETQNTLVKQLQALQONSSVLSDEKSKRLTIL 120
 DB 61 NNNNAGDKWSAFLEKQSTLAQMPLOETQNTLVKQLQALQONSSVLSDEKSKRLTIL 120
 QY 121 NTMSTIYSTGKVCNPDNPOECILLEPGINEIMANSLDYNERTLMAESMRSEVGKQLRPLY 180
 DB 121 NTMSTIYSTGKVCNPDNPOECILLEPGINEIMANSLDYNERTLMAESMRSEVGKQLRPLY 180
 QY 181 EEEVVLKEMARAHNYEDYGYRWGDIYEVNGVDGYDSRQQLIEDVEHTEEIKPLYEHL 240
 DB 181 EEEVVLKEMARAHNYEDYGYRWGDIYEVNGVDGYDSRQQLIEDVEHTEEIKPLYEHL 240
 QY 241 HAYYRAKIMNAPPSYISPIGCLPAHLGDMWGRFMTLYSLTFPFGOKPVIDVTDAMVDQ 300
 DB 241 HAYYRAKIMNAPPSYISPIGCLPAHLGDMWGRFMTLYSLTFPFGOKPVIDVTDAMVDQ 300
 QY 301 AMDQRIFFKAKEFFVSGLEPNMTQGWENSMLTDPGNVOKAVCHPTAMDIGKDFRILM 360
 DB 301 AMDQRIFFKAKEFFVSGLEPNMTQGWENSMLTDPGNVOKAVCHPTAMDIGKDFRILM 360
 QY 361 CTKVTMDFLTAHHEMHIQYDMAAQAOPFLNRNAGEGFHEAVGETMISLATAPRHLKS 420
 DB 361 CTKVTMDFLTAHHEMHIQYDMAAQAOPFLNRNAGEGFHEAVGETMISLATAPRHLKS 420
 QY 421 IGLISPDFOEDNTEINFLKQALITVGLPFTYMLEKRWMMVFKGEIPKQMMKKWEM 480
 DB 421 IGLISPDFOEDNTEINFLKQALITVGLPFTYMLEKRWMMVFKGEIPKQMMKKWEM 480
 QY 481 KREIVGVPEVPDHEMCDPASLFHVSNDYSEFRITRTILYQFOFODALCOAKHEGPLH 540
 DB 481 KREIVGVPEVPDHEMCDPASLFHVSNDYSEFRITRTILYQFOFODALCOAKHEGPLH 540
 QY 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVGAKNNVRLNLYFPPLFTWLKDDOK 600
 DB 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVGAKNNVRLNLYFPPLFTWLKDDOK 600
 QY 601 NSFVGMSTWSPYADOSIKVRISLKSALGDKAYEMNDNEMYLFRSSVAYAMROYFLKVN 660
 DB 601 NSFVGMSTWSPYADOSIKVRISLKSALGDKAYEMNDNEMYLFRSSVAYAMROYFLKVN 660
 QY 661 QMILFGEEDVRVANKPRISFNFVTAAPKNVSDIIPREVEKAIKRSRINDARLNDN 720
 DB 661 QMILFGEEDVRVANKPRISFNFVTAAPKNVSDIIPREVEKAIKRSRINDARLNDN 720
 QY 721 SLEFLGIQPTLGPNNPVSIMLIVGVGVVIGVILITFTGIRDRKKKKKASGENP 780
 DB 721 SLEFLGIQPTLGPNNPVSIMLIVGVGVVIGVILITFTGIRDRKKKKKASGENP 780
 QY 781 YASIDISKGNNPGFQNTDDVQTSF 805
 DB 781 YASIDISKGNNPGFQNTDDVQTSF 805

RESULT 7
 AAU09092
 ID AAU09092 standard; Protein; 711 AA;
 AC AAU09092;
 XX
 XX 20-DEC-2001 (first entry)
 DE Novel human protein NHP #1.
 XX

Human: novel human protein: NHP; antidiabetic; antirheumatic;
 antirheumatic; cytosolic; antiarteriosclerotic; vulnerary;
 neuroprotective; neurotropic; antiparkinsonian;
 anti-human immunodeficiency virus; antistomatitic; vasotropic; cardiant;
 hypotensive; anorectic; antileptilitic; neuroleptic; anticonvulsant;
 antimalarial; immunosuppressive; cerebroprotective; antimicrobial;
 antileptilitic; antibacterial; antiparasitic; thyromimetic;
 immunomodulatory; antiseborrheic; dermatological; vasoconstriction;
 gastrointestinal disorder; cardiovascular disorder; hypertension;
 coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;
 coxemia; male infertility; impotence; testicular cancer; lung tumour;
 hyperproliferative disorder; pulmonary system disorder;
 central nervous system disorder; bone disorder; Parkinson's disease;
 neurodegenerative disease; Alzheimer's disease; dementia; paranoia;
 Huntington's disease; schizophrenia; amyotrophic lateral sclerosis;
 panic disorder; learning disabilities; immune system disorder;
 psychosis; autism; sleep disorder; immune system disorder;
 Hashimoto's thyroiditis; musculo-skeletal system disorders;
 multiple sclerosis; ischemic brain injury; stroke; infectious disease;
 diabetes mellitus; immunological disorder; asthma; AIDS; immunogen;
 acquired immunodeficient syndrome; leukaemia; rheumatoid arthritis;
 inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;
 neural system disorder; respiratory disorder; olfactory disorder;
 wound healing; chromosome X.

Homo sapiens.

Key	Location/Qualifiers
Domain	1..681 Extracellular_domain
Region	48..55
Region	/label= Immunogenic_epitope
Region	110..118
Region	/label= Immunogenic_epitope
Region	136..146
Region	/label= Immunogenic_epitope
Region	151..158
Region	/label= Immunogenic_epitope
Misc-difference	219
FT	/label= OTHER
FT	/note= "Other- Any amino acid encoded by WST"
FT	240
FT	/label= OTHER
FT	/note= "Other- Any amino acid encoded by RCC"
FT	499
FT	/label= OTHER
FT	/note= "Other- Any amino acid encoded by NTT"
FT	682..698
FT	/label= Transmembrane_domain
XX	
PN	WO200174896-A1.
PD	11-OCT-2001.
PE	02-APR-2001: 2001WO-US10542.
PR	03-APR-2000: 2000US-194118P.
PR	29-SEP-2000: 2000US-236384P.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PI	Moore PA, Ni J, Soppet DR, Coleman TA, Gentz RL, Endress GA;
PI	LI Y, Dillon PJ;
DR	WPI: 2001-626394/72.
DR	N-PSDB; AAS14880.
XX	
PT	New human proteins, useful for diagnosing, treating, preventing and/or
PT	prognosing disorders related to the proteins, including cardiovascular
PT	disorders, autoimmune disorders and reproductive disorders
XX	
PS	Claim 11, Page 298-301; 318pp; English.
XX	

CC The invention relates to novel human proteins (NHP) and the
 CC nucleic acids that encode them and antibodies raised against them.
 CC The proteins, antibodies and nucleic acids are useful in the diagnosis,
 CC prognosis, prevention and/or treatment of diseases and/or disorders
 CC involving vasoconstriction, gastrointestinal disorders, cardiovascular
 CC disorders (e.g. hypertension, erectile dysfunction, high blood pressure,
 CC coronary heart disease and arteriosclerosis), anorexia, obesity, bulimia,
 CC coxemia, disorders of small intestine, disorders of reproductive system
 CC (e.g. male infertility and/or impotence), testicular cancer, lung tumours
 CC and other hyperproliferative disorders, bone disorders, neurodegenerative
 CC central nervous system disorders, bone disorders, Parkinson's
 CC disease, Huntington's disease, schizophrenia, mania, dementia, paranoia,
 CC panic disorder, learning disabilities, amyotrophic lateral sclerosis,
 CC psychoses, autism, sleep disorders, immune system disorders (e.g.
 CC Hashimoto's thyroiditis), renal and musculo-skeletal system disorders,
 CC central nervous system disorders (e.g. multiple sclerosis, ischemic
 CC brain injury and/or stroke), infectious diseases, diabetes mellitus,
 CC immunological disorders (e.g. asthma, acquired immunodeficient syndrome
 CC (AIDS), leukaemia, rheumatoid arthritis, inflammatory bowel disease,
 CC sepsis, acne, psoriasis and lupus erythematosus), neural system
 CC disorders, respiratory disorders, olfactory disorders and wound
 CC healing. The present sequence represents an NHP of the invention the
 CC gene for which is located on the X chromosome.

Sequence 711 AA;

Query Match 88.04; Score 3775; DB 22; Length 711;
 Best Local Similarity 99.35; Pred. No. 1.3e-311; 32X
 Matches 704; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY	62	MNNAGDKWSAFLEKREOSTLAQWYPIQELQNLVVKLQALQALQONSSVLSSEDKSRKLTNTIN	121
DB	1	MNNAGDKWSAFLEKREOSTLAQWYPIQELQNLVVKLQALQALQONSSVLSSEDKSRKLTNTIN	60
QY	122	TWSTYTGKVCNPNDECELLLEBGLNETMANSIDYNERLMAESRSPGOLRPLYE	181
DB	61	TWSTYTGKVCNPNDECELLLEBGLNETMANSIDYNERLMAESRSPGOLRPLYE	120
QY	182	EVVVLKNEKARANHEDYGDYWRGDEYVNGVDYDYSRGQLIEDVETFEIRPLEYELH	241
DB	121	EVVVLKNEKARANHEDYGDYWRGDEYVNGVDYDYSRGQLIEDVETFEIRPLEYELH	180
QY	242	AVYRAKLNAYSPYISPIGCLPAHLGDMGGRFWNTLYSLTPPGQKPNIDYDAMVDA	301
DB	181	AVYRAKLNAYSPYISPIGCLPAHLGDMGGRFWNTLYSLTPPGQKPNIDYDAMVDA	240
QY	302	WDAQRIFKAEKFFVSVGLPNTGOFWENSMITDGGANYQKAVCHPTAMDLSKGFRTLMC	361
DB	241	WDAQRIFKAEKFFVSVGLPNTGOFWENSMITDGGANYQKAVCHPTAMDLSKGFRTLMC	300
QY	362	TKVTNDPFLAHHEMGIQDMAVAOPFLFNGANGEGFHEAVGEIMSLSATPKHLKSI	421
DB	301	TKVTNDPFLAHHEMGIQDMAVAOPFLFNGANGEGFHEAVGEIMSLSATPKHLKSI	360
QY	422	GLSPDFQEDNETEINFLKQALITVGTLPFTYMLEKRMWYFKETPKDDMMKKWEMK	481
DB	361	GLSPDFQEDNETEINFLKQALITVGTLPFTYMLEKRMWYFKETPKDDMMKKWEMK	420
QY	482	REIVGAVEPVPHEVTCDPASLFHVSNDYSFRTYTRTYQFOFOEALCOAKHEGRLHK	541
DB	421	REIVGAVEPVPHEVTCDPASLFHVSNDYSFRTYTRTYQFOFOEALCOAKHEGRLHK	480
QY	542	CDISNSTEAGOKLFNMLRKSEPTLALENVYGAKNMVRPLNYFEPDLTWLKDOKN	601
DB	481	CDISNSTEAGOKLFNMLRKSEPTLALENVYGAKNMVRPLNYFEPDLTWLKDOKN	540
QY	602	SFVGSSTDSWSPYDQSTKVAISLSKALGKAVEMNDENYLRSSVAYAMROEFLKVNQ	661
DB	541	SFVGSSTDSWSPYDQSTKVAISLSKALGKAVEMNDENYLRSSVAYAMROEFLKVNQ	600
QY	662	MILFGEEDVAVANLKPRISFNFTVAPKVVSDIIPTEVEKAIKRSRSLINDAPLANS	721

The invention relates to the metalloproteinase *zace2*, *zace2*, an angiotensin-converting enzyme is a zinc metalloproteinase that plays roles in blood pressure regulation and fertility. *zace2* can be expressed by standard recombinant methodology. *zace2* polypeptides are useful for treating an inflammatory bowel disease. *zace2* polypeptides are useful for treating a disease associated with inflammation like arthritis and ulcerative colitis) as targets for identifying modulators of zinc protease activity, for screening or identifying new angiotensin-converting enzyme (ACE) inhibitors, and as a basis for rational drug design for inhibitory molecules. The nucleic acids can be used to detect the expression of a

```

|:| | | | | | | | | |
781 YDSMPTCECEENACTOVEDD'CEET CCE
b

```

Inhibitors of ACE are used for treating hypertension of various

QY
1 M S S S W L L S L V A V I A Q S T I E E Q A N I F E D N I N H E B D Z I G C O R N

AC MAU12207;

XX 24-OCT-2001 (first entry)
 DT
 XX
 DE Human PRO185 polypeptide sequence.
 XX
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 KW adipocyte; A-peptide; factor VIIa; gene therapy.
 OS
 XX Homo sapiens.
 XX
 PN WO200140466-A2.
 XX
 PD 07-JUN-2001.
 XX
 PF 01-DEC-2000; 2000WO-US32678.
 XX
 PR 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 PR 09-DEC-1999; 99US-0170262.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30911.
 PR 20-DEC-1999; 99WO-US30999.
 PR 30-DEC-1999; 99WO-US31243.
 PR 06-JAN-2000; 2000WO-US00277.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US03365.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 24-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US04914.
 PR 24-FEB-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 21-MAR-2000; 2000WO-US07532.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 10-NOV-2000; 2000WO-US30873.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Geritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX
 DR WPI; 2001-408281/43.
 DR N-PSDB; AAS21279.
 XX
 XX
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical
 XX
 PS Claim 12; Fig 72; 813pp; English.
 XX
 XX AAU12172-AAU12446 represent novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from

CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, or the release of a cytokine from peripheral blood
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 XX
 SQ Sequence 555 AA;

Query Match 69.4%; Score 2979; DB 22; Length 555;
 Best Local Similarity 99.8%; Pred. No. 3.8e-244;
 Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSWLLSLVAVTAOSTIEBOAKTFLDKFHEAEDLFYQSSLSAMWNTNTEENVQ 60
 DB 1 MSSSSWLLSLVAVTAOSTIEBOAKTFLDKFHEAEDLFYQSSLSAMWNTNTEENVQ 60
 QY 61 MNNAAGDKMSAFLEKOSTLQMYPLEIQNTLYKLDLQONGSSVLEDSKRLNTIL 120
 DB 61 MNNAAGDKMSAFLEKOSTLQMYPLEIQNTLYKLDLQONGSSVLEDSKRLNTIL 120
 QY 121 NTMSTIYSTGKVCNPNPQECILLEPGLNEIMANSIDYNERLWAMSMRSEVGKOLRPLY 180
 DB 121 NTMSTIYSTGKVCNPNPQECILLEPGLNEIMANSIDYNERLWAMSMRSEVGKOLRPLY 180
 QY 181 EEEVVLKNEMARANHYEDGDYRGDYEVNGDGYDSRGQLEIEVHTFEIKPLYEHL 240
 DB 181 EEEVVLKNEMARANHYEDGDYRGDYEVNGDGYDSRGQLEIEVHTFEIKPLYEHL 240
 QY 241 HAVVRAKLMAVPSYISPIGCLPAHLGDMGREFNTNLSLTPRGCKPNTIDVDAWDQ 300
 DB 241 HAVVRAKLMAVPSYISPIGCLPAHLGDMGREFNTNLSLTPRGCKPNTIDVDAWDQ 300
 QY 301 AMDAQRIFKEAEKFFVSVGLPMTQGEWNSMLTDPGVORAVCPHTAMDGKGFRTILM 360
 DB 301 AMDAQRIFKEAEKFFVSVGLPMTQGEWNSMLTDPGVORAVCPHTAMDGKGFRTILM 360
 QY 361 CTKVTMDDELTAHHEMGHIOYDMAVAAOPLIRNGANGFEHANGELMSAATPKILKS 420
 DB 361 CTKVTMDDELTAHHEMGHIOYDMAVAAOPLIRNGANGFEHANGELMSAATPKILKS 420
 QY 421 IGLSPDFQEDNETINLTKQALTYGTLPTMYLEKRWVFKGEIPKDDMMKKWMEK 480
 DB 421 IGLSPDFQEDNETINLTKQALTYGTLPTMYLEKRWVFKGEIPKDDMMKKWMEK 480
 QY 481 KREIVGVEPVPHDETYCDPASLFHVSNDYSFIYRTLYXQFQFALCQAAKHGEPILH 540
 DB 481 KREIVGVEPVPHDETYCDPASLFHVSNDYSFIYRTLYXQFQFALCQAAKHGEPILH 540
 QY 541 KCDISNSTEAGOKL 554
 DB 541 KCDISNSTEAGOKL 554

RESULT 11
 AAU67311
 ID AAU67311 standard; protein; 480 AA.
 XX
 AC AAU67311;
 XX
 DT 11-APR-2000 (first entry)
 XX
 DE Human MPRO15 amino acid sequence #2.
 XX
 KW MPRO15; treatment; hypertension; human; myocardial disease; apoplexy;
 KW heart disease; apoplexy; heart disease; nervous denaturation; hormone;
 KW Alzheimer's disease; cytokine.
 XX
 OS Homo sapiens.

PN JP11318472-A.
 KW 24-NOV-1999.
 XX 22-JAN-1999; 99GP-0014949.
 XX 13-MAY-1998; 98GB-0010373.
 PR 18-AUG-1998; 98GB-0018009.
 XX (SMIK) SMITHKLINE BEECHAM PLC.
 PA WPI: 2000-109268/10.
 DR MPROT15 polypeptide and MPROT15 polynucleotides - useful for the
 XX treatment of hypertension, myocardial diseases, apoplexy, heart
 PT diseases, nervous denaturation, Alzheimer's disease etc.
 PS Claim 19; Page 20-21; 22pp; Japanese.
 CC This is amino acid sequence #2 of human MPROT15. The MPROT15
 CC polynucleotide and polypeptide sequences can be used for the treatment of
 CC hypertension, myocardial diseases, apoplexy, heart diseases, nervous
 CC denaturation, Alzheimer's disease and diseases related to the processing
 CC of peptide hormones and cytokines.
 CC
 XX Sequence 480 AA:
 SQ
 Query Match 59.2%; Score 2539; DB 21; Length 480;
 Best local similarity 100.0%; Pred. No. 7.4e-207;
 Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 LVAVTAAGSTIEHQAFTFDKFNHAEADLFYQSSLSASWNTNTTEENVQNMNAGDKMS 70
 Db 10 LVAVTAAGSTIEHQAFTFDKFNHAEADLFYQSSLSASWNTNTTEENVQNMNAGDKMS 69
 QY 71 AFLKEQSTLAQMTPLQELQNTLVKLOLQLOALQONGSSVSEDESKRLNTLTMSTIYSTG 130
 Db 70 AFLKEQSTLAQMTPLQELQNTLVKLOLQLOALQONGSSVSEDESKRLNTLTMSTIYSTG 129
 QY 131 KVCMPDNPQECLEPELGNLIMANSIDYNERLWAMESRSEVQKOLRPLYEYEVYLNKEM 190
 Db 130 KVCMPDNPQECLEPELGNLIMANSIDYNERLWAMESRSEVQKOLRPLYEYEVYLNKEM 189
 QY 191 ARANHEDEYDGYWNGDEYVNGVDGYDSRQQLIEDVEHFEETKPLYEHLHAYVRAKLMN 250
 Db 190 ARANHEDEYDGYWNGDEYVNGVDGYDSRQQLIEDVEHFEETKPLYEHLHAYVRAKLMN 249
 QY 251 AVPTSTPITGCLPAHLIGDMWGRFTNLVSLTVPFGOKPNIDVTAMDADQRIKE 310
 Db 250 AVPTSTPITGCLPAHLIGDMWGRFTNLVSLTVPFGOKPNIDVTAMDADQRIKE 309
 QY 311 AEKFEVSQGLPMTQGFENSMETDPGNVOKAVCHPTANLQGGDRILLMCTKVTMDPDL 370
 Db 310 AEKFEVSQGLPMTQGFENSMETDPGNVOKAVCHPTANLQGGDRILLMCTKVTMDPDL 369
 QY 371 TAAHEHGHTQYDMAYAAQPELLRNANGCFHEAVGEIMSSAATPKHLKSGILSPDFGE 430
 Db 370 TAAHEHGHTQYDMAYAAQPELLRNANGCFHEAVGEIMSSAATPKHLKSGILSPDFGE 429
 QY 431 DNTEINFLKQALITVGLTPTTYMLEKRWMTVFGELTPKQOMKMKWEMK 481
 Db 430 DNTEINFLKQALITVGLTPTTYMLEKRWMTVFGELTPKQOMKMKWEMK 480
 RESULT 12
 ID AA009102 standard; Protein; 261 AA.
 XX AA009102;
 AC 20-DEC-2001 (first entry)
 DT
 XX
 DE Novel human protein NHP #11.

XX Human; novel human protein; NHP; antidiabetic; antineumatic;
 KW antiarthritic; cytostatic; antiarteriosclerotic; vulnerary;
 KW neuroprotective; nootropic; antiparkinsonian;
 KW anti-human immunodeficiency virus; antiautism; vasotropic; cardiatic;
 KW hypotensive; anorectic; antilethargy; neuroleptic; anticonvulsant;
 KW antitumor; immunosuppressive; cerebroprotective; antimicrobial;
 KW immunomodulatory; antibacterial; antiparasitic; thyromimetic;
 KW immunomodulatory; antiseborrheic; dermatological; vasoconstriction;
 KW gastrointestinal disorder; cardiovascular disorder; hypertension;
 KW coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;
 KW cachexia; male infertility; impotence; testicular cancer; lung tumour;
 KW hyperproliferative disorder; pulmonary system disorder;
 KW central nervous system disorder; bone disorder;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW Huntington's disease; schizophrenia; mania; dementia; paraneoplasia;
 KW panic disorder; learning disability; immune system disorder;
 KW psychosias; autism; sleep disorder; immune system disorder;
 KW Hashimoto's thyroiditis; musculo-skeletal system disorders;
 KW multiple sclerosis; ischemic brain injury; stroke; infectious disease;
 KW diabetes mellitus; immunological disorder; asthma; AIDS; immunogen;
 KW acquired immunodeficient syndrome; leukaemia; rheumatoid arthritis;
 KW inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;
 KW neural system disorder; respiratory disorder; olfactory disorder;
 KW wound healing.
 XX
 OS Homo sapiens.
 XX
 XX WO2001/74896-A1.
 XX 11-OCT-2001.
 XX 02-APR-2001; 2001WO-US10542.
 XX 03-APR-2000; 2000US-194118P.
 XX 29-SEP-2000; 2000US-236384P.
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX Moore PA, Ni J, Soppet DR, Coleman TA, Gentz RL, Endress GA;
 PI Li Y, Dillon PJ;
 PI WPI: 2001-626394/72.
 DR N-PDSB; AAS14890.
 DR
 XX New human proteins, useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cardiovascular
 PT disorders, autoimmune disorders and reproductive disorders -
 XX
 PS Claim 11; Page 311-312; 318pp; English.
 XX The invention relates to novel human proteins (NHP) and the
 XX nucleic acids that encode them and antibodies raised against them.
 CC The proteins, antibodies and nucleic acids are useful in the diagnosis,
 CC prognosis, prevention and/or treatment of diseases and/or disorders
 CC involving vasoconstriction, gastrointestinal disorders, cardiovascular
 CC disorders (e.g. hypertension, erectile dysfunction, high blood pressure,
 CC coronary heart disease and arteriosclerosis), anorexia, obesity, bulimia,
 CC cachexia, disorders of small intestine, disorders of reproductive system
 CC (e.g. male infertility and/or impotence), testicular cancer, lung tumours
 CC and other hyperproliferative disorders, disorders of pulmonary system,
 CC central nervous system disorders, bone disorders, neurodegenerative
 CC diseases and behavioural disorders (e.g. Alzheimer's disease, Parkinson's
 CC disease, Huntington's disease, schizophrenia, mania, dementia, paraneoplasia,
 CC panic disorder, learning disability), immune system disorders (e.g.
 CC psychosias, autism, sleep disorders), infectious diseases, ischaemic
 CC Hashimoto's thyroiditis), renal and musculo-skeletal system disorders,
 CC brain injury and/or stroke), infectious diseases, diabetes mellitus,
 CC immunological disorders (e.g. asthma, acquired immunodeficient syndrome
 CC (AIDS), leukaemia, rheumatoid arthritis, inflammatory bowel disease,
 CC sepsis, acne, psoriasis and lupus erythematosus), neural system
 CC disorders, respiratory disorders, olfactory disorders and wound

CC healing. The present sequence represents an NHP of the invention.
 XX
 SQ Sequence 261 AA;

Query Match 31.7%; Score 1359; DB 22; Length 261;
 Best Local Similarity 99.6%; Pred. No. 5.3e-107;
 Matches 252; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 62 MNNAGDKWSAFLEKQSTLAQMPLOEIQNLTVKLOLQALQNGSSVLESDSKRLNTILN 121
 DB 1 MNNAGDKWSAFLEKQSTLAQMPLOEIQNLTVKLOLQALQNGSSVLESDSKRLNTILN 60
 QY 122 TMSITSTGKVCNPDNDPQECLEPELNEIMANSIDYNERLAMESRSEYKQRLPYE 181
 DB 61 TMSITSTGKVCNPDNDPQECLEPELNEIMANSIDYNERLAMESRSEYKQRLPYE 180
 QY 182 EYVVLKNEARAHNHYEDYDGYRSDYEVNGVDGYDYSKGLIEDEHTEFEIKRPLYEHLH 241
 DB 121 EYVVLKNEARAHNHYEDYDGYRSDYEVNGVDGYDYSKGLIEDEHTEFEIKRPLYEHLH 180
 QY 242 AYVRKALMAYPSYISPIGCLPAHLGDMGREFWNTLSLTVPGQKPNIDVTDAWVDA 301
 DB 181 AYVRKALMAYPSYISPIGCLPAHLGDMGREFWNTLSLTVPGQKPNIDVTDAWVDA 240
 QY 302 WDAORIFKEAEKF 314
 DB 241 WDAORIFKEAEKF 253

RESULT 13

AA010426
 ID AAR10426 standard; Protein; 732 AA.
 XX
 AC AAR10426;

DT 10-APR-1991 (first entry)
 XX

DE Human testicular angiotensin conversion enzyme.
 KW human testicular angiotensin conversion enzyme; TACE;
 KM male sterility.
 XX Homo sapiens.
 OS

XX
 FH Key

FT Peptide Location/Qualifiers
 FT 1..21
 FT /label= signal peptide
 FT 22..732
 FT Protein /label= mature TACE

PN W09100354-A.
 XX
 PD 10-JAN-1991.
 XX
 PE 05-JUL-1990; 90WO-FR00513.
 XX
 PR 05-JUL-1989; 89FR-0009062.
 XX
 PA (INRM) INST NAT SANTE RECH.
 XX
 PI Soubrier F, Alhenc-Gelas F, Hubert C, Corvol P,
 XX WPI, 1991-036748/05.
 DR N-PSDB; AAO10328.
 DR
 XX
 PT Nucleic acid - encoding human testicular angiotensin conversion
 PT enzyme, used e.g. for in vitro detection of enzyme in organism
 XX
 PS Claim 1; Fig 1; 48pp; French.
 XX
 CC A bank of human testicular cDNA in Lambda gt11 was screened with a
 CC complete sequence of TACE was reconstructed from 4 separate clones.

CC The isolated nucleic acid sequence was inserted into a plasmid for
 CC expression of the protein. The invention covers polypeptides
 CC containing all or part of TACE sequence. These are useful in
 CC treatment of inflammation or infectious diseases, especially acute
 CC pancreatitis, or diseases in which kinins are involved. Antibodies
 CC against the polypeptides are useful as immunoassay reagents for
 CC TACE.
 XX
 SQ Sequence 732 AA;

Query Match 31.3%; Score 1344; DB 12; Length 732;
 Best Local Similarity 41.8%; Pred. No. 5e-105;
 Matches 259; Conservative 119; Mismatches 204; Indels 38; Gaps 10;

QY 15 TAAOS-----TTEQANTFLDKFNEHAEDELTYQSSLSANNYNNTTEE-----NVQNM 62
 DB 61 TSAQSPVLVDEAEASKFVEEDRTSQVYWMNEYAAMNNTNTTTSKILLQKNMOIA 120
 QY 63 NNAQDKWSAFLEKQSTLAQMPLOEIQNLTVKLOLQALQNGSSVLESDSKRLNTILN 120
 DB 121 NHT-----LKGTQAKKFDVNOLOMTTKRIKRVQDLERALPAQLEENKILLD 172
 QY 123 MSTIYSTGKVCNPDNDPQECLEPELNEIMANSIDYNERLAMESRSEYKQRLPYE 182
 DB 173 MSTIYSTGKVCNPDNDPQECLEPELNEIMANSIDYNERLAMESRSEYKQRLPYE 180
 QY 183 YVVLKNEARAHNHYEDYDGYRSDYEVNGVDGYDYSKGLIEDEHTEFEIKRPLYEHLH 242
 DB 231 YVELINQAAARUNGVDYDAGDSKRSMTETPSLE-----QDLERLFOELQPLILNHA 280
 QY 243 AYVRKALMAY-PSYISPIGCLPAHLGDMGREFWNTLSLTVPGQKPNIDVTDAWVDA 301
 DB 281 YVRALRHHYGAOHINLEGPVPAHLGDMGREFWNTLSLTVPGQKPNIDVTDAWVDA 240
 QY 302 WDAORIFKEAEKFVSGLPNMOCGWENSMILDPGNQKAVCHPTANDLKG--DEFRLM 360
 DB 341 WTPRMFKKADDFPTSLGLLPVPEFWMKSMLEKPTDGEVYCHASAMDFYNGKDFIKQ 400
 QY 361 CTKYTMDFLTANHEGHIQYDWAYAAOPLFLRANGANGFEHAGVETMSAATPKHLKS 420
 DB 401 CTYVNEEDLVAAHEHGHIOYFMQYKDLVYALREBANGFEHAGVETMSAATPKHLKS 420
 QY 421 IGLISPDFOEDNETETINFLKQALTITVGTPTWLEKRWMEKGEIPDMMKKWME 480
 DB 461 IMLISSEGGSD--EHDINFLMKALDKIAIFPFSYLVQMRVRFVDSITKENYNOEMNSI 519
 QY 481 KREIVGVAPVPVHDETYCDPASLEFVNSDYSFIRYTRTYLXOFQFOALCOAKHEGPIH 540
 DB 520 RLKYGQLCPVYRTQGDGDPGAKFHIPSSVYIRYFSTIIOFHQFHALCOAGHGTGPIH 579
 QY 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVYGAKNMNVRLNYPFPLTWLKDQK 600
 DB 580 KCDIYQSKKQGRALATAMKLCFSRPEAMOLITGONMSASAMLSYKFLDWMRTENE 639
 QY 601 --NSFVGW-STDMSPYADQS 617
 DB 640 LHGEKLGMPQYNTWTPNSARS 659

RESULT 14

AA020501
 ID AAO20501 standard; Protein; 1265 AA.
 XX
 AC AAO20501;

DT 27-JUN-2002 (first entry)
 XX
 DE Protein of APP related human homologue hcp51674.
 KW Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;
 KW amyloid precursor protein; tissue-specific expression control; human APP;
 KW APP pathway modulator; gene therapy.
 XX

FT Modified-site 1191, 1193
 FT /label-putative N-glycosylation site
 FT Modified-site 1225, 1227
 FT /label-putative N-glycosylation site
 PN WO9003435-A.
 PD 05-APR-1990.
 XX
 XX 27-SEP-1989; 89WO-FR00496.
 XX
 XX 27-SEP-1989; 88FR-0012620.
 XX
 PA (INRM) INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE.
 PI Soudrier F, Alhenc-Gelas F, Hubert C, Corvol P;
 DR WPI: 1990-132272/17.
 DR N-PSDB; AA004027.
 XX
 PT New DNA encoding human angiotensin converting enzyme used eg in
 PT diagnosis of hypertension, evaluation of enzyme inhibitors
 PS
 PS Disclosure; ; p; French.
 CC Human angiotensin converting enzyme hydrolyses angiotensin I and kinins.
 CC Either intact enzyme or fragments thereof can be used to generate
 CC antibodies for diagnostic use. Oligonucleotide probes can also be made
 CC which are complementary to the sequence encoding the enzyme.
 XX
 SQ Sequence 1306 AA;
 Query Match 31.2%; Score 1337; DB 11; Length 1306;
 Best Local Similarity 41.7%; Pred. No. 4.8e-104;
 Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;
 QY 20 TTEBQAKTFLDKFNHAEADLYGSSLASWNTNITEE-----NVCNMNAGDKMSA 71
 DB 644 TTEAKSKFVEEYDTSQVWNEEYAEANMNYNTITETSKILLQNMJAHNT----- 697
 QY 72 FLKEOSTLAQYPILOETIOLNFKLOALQNGSSVSEKSKRLTITNTSTYSGK 131
 DB 698 --LKGXTQARKFDVNOQNTIKRIKKVODLERALPQOELEENKQIILDETTYSVAT 755
 QY 132 VCNPDNPECLLEPGELNEMANSIDYNERLWAMESWSEVCKOIRPYEEVYVKNMA 191
 DB 756 VCHNG--SCLQEPDLTNVMTSRKYEDLLWAMEGRDKAGRAILOFPYKVELINQAA 813
 QY 192 RANHIEDYGDYWRGDEYVNGVDYDSRGQLLEDVHENTFEIKPLYEHLAAVYRAKLMA 251
 DB 814 RLNGYVDAGDSWRSKMYETPSLE-----QDLERLFQELQPLYLMLAAVYRRALHRH 863
 QY 252 Y-PSYISPIGCLPAHLGLDMGRFNTNLYSLTVPGRPNIDVIDAMVDQMDAQRERKE 310
 DB 864 YGAQHINLEGPJPAHLGLMMAQMTSNITDYLVPPSPASMDTTEANLKGWPRRKEFE 923
 QY 311 AEKFEVSVGLPMTQGFWENSMULTDGNVOKAVCHPTAMDIGKC-DRITIMCTKVTWDF 369
 DB 924 ADDEFFSLILPVPPEFNKSKLEKPTGREGVCHASAMQFENCKDRICQCTVNLDEL 983
 QY 370 LTAHMHGHIQYDAAIAAOPFLLRNANGNEGFHEAVGEMISLSTAPRHLKSLISPDQ 429
 DB 984 VVAHHEMGIQYFQYKDLPAALREGANPGFHEAIGDYALSTVRHLHSLMLSSSEG 1043
 QY 430 EDNETEINFLKQALTIYGLTLPYMLKRWKMYFKEIKDKQMKRWEMKREIYGYE 489
 DB 1044 SD-EHDINFLMKMLDKAIFESYLDVDMRWRFEDGSLTKENYNDEMWSLRLKYGICLP 1102
 QY 490 PVPHEETCDPASLFEHVSNDYSFTRYYTTLIOFOFQALCOAKKEGSLCDISNSTE 549
 DB 1103 FVPRKQGFDFGAKFHLPSYPYIRYFVFTIQFQHEALCOAGHTGPHLHCDDIYOSKE 1162
 QY 550 AGORLFLNMLRLGKSEPTLALENVGAKMMNVRPLNLFEPFLTWLKDONK--NSFYGV- 606

DB 1163 AGORLATAMKLGSRPWPAPMOLITGQPNMSASAMLSYFKPLDMLRTENELHGEKLGWP 1222
 QY 607 STDWSPYADQS 617
 DB 1223 QYNWTPNSARS 1233
 RESULT 16
 AAM68155
 ID AAM68155 standard; Protein; 1306 AA.
 XX
 AC AAM68155;
 XX
 DT 09-NOV-1998 (first entry)
 XX
 DE Human angiotensin converting enzyme.
 XX
 KW Angiotensin converting enzyme; ACE; hypertension; exercise; human;
 XX genetic marker.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..29
 FT Protein 30..1307
 FT /label= Sig-peptide
 FT /label= Mat-protein
 XX
 PN WO9831835-A1.
 XX
 PD 23-JUL-1998.
 XX
 PF 22-DEC-1997; 97WO-US22974.
 PR 27-MAY-1997; 97US-0048309.
 PR 16-JAN-1997; 97US-0035382.
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 PA (UYPL-) UNIV PITTSBURGH.
 PI Ferrell RE, Hagberg JW;
 DR WPI: 1998-414128/35.
 DR N-PSDB; AAV41320.
 XX
 PT Analysis of genetic markers to identify subjects who will benefit
 PT from exercise - also assessing risk of cardiovascular disease from
 PS angiotensin-converting enzyme genotype
 PS Disclosure; Page 35-41; 61pp; English.
 PS
 XX This is human angiotensin converting enzyme (ACE). The ACE gene
 XX (see AAV41320) is polymorphic with 2 common alleles (I and D),
 XX resulting in 3 genotypes, II, ID and DD. It is an object of the
 XX invention to identify individuals possessing a certain genotype and
 XX associated ailment, and to determine if the health of that
 XX individual can be improved by altering behavior. A claimed method
 XX comprises identifying individuals having a certain phenotype and
 XX determining the presence or absence of genetic markers associated
 XX with the phenotype, and instituting a lifestyle change to exploit
 XX or counteract the phenotype expressed by the gene marker. If the
 XX phenotype is hypertension, the gene marker is at least one
 XX insertion (I) ACE allele and diastolic blood pressure. The gene marker
 XX can be identified by PCR amplification (see AAV41321-22) of the
 XX appropriate gene region. The general method can be used to
 XX identify subjects who will benefit most from physical exercise
 XX and also to identify those who are likely to be successful in
 SQ Sequence 1306 AA;

11-JUL-2000; 2000US-0614150.

[illegible]

DB	597	NIKNNVHIGWTT	608
RESULT	20		
AAR70013			
ID	AAR70013	standard; Protein; 660 AA.	
XX			
AC	AAR70013;		
XX			
DT	25-SEP-1995	(first entry)	
XX			
DE	Tick carboxypeptidase.		
XX			
KM	Tick; antigen; carboxypeptidase; vaccine.		
XX			
OS	Boophilus microplus.		
XX			
FH	Key	Location/Qualifiers	
FT	Peptide	30..53	
FT		/label= BM91 peptide	
FT	Peptide	185..201	
FT		/label= peptide AAT91251	
FT	Peptide	202..209	
FT		/label= AAT91141	
FT	Peptide	225..253	
FT		/label= T9118	
FT	Peptide	262..269	
FT		/label= BM91 peptide	
FT	Peptide	442..452	
FT		/label= T9129(a)	
FT	Peptide	456..481	
FT		/label= T9129(b)	
FT	Peptide	599..606	
FT		/label= T9109	
FT	Peptide	1..29	
FT		/label= signal	
FT	Domain	639..655	
FT		/label= C-terminal transmembrane domain	
FT		/note= "putative"	
XX			
PN	WO9504827-A.		
XX			
PD	16-FEB-1995.		
XX			
PF	10-AUG-1994;	94WO-AU00463.	
XX			
PR	10-AUG-1993;	93AU-0000458.	
XX			
PA	(BIOT-) BIOTECH AUSTRALIA PTY LTD.		
XX	(CSIR) COMMONWEALTH SCI & IND RES ORG.		
XX			
PI	Codon GS, Kemp DH, Tellam RL, Willadsen P;		
XX			
DR	WPI; 1995-090905/12.		
XX			
DR	N-PSDB; AAQ82948.		
XX			
PT	New antigenic tick carboxypeptidase and corresp. DNA - are used		
PT	in vaccines for producing antibodies against ticks, insects and		
XX	nematodes.		
XX			
PS	Disclosure: Fig 6; 138pp; English.		
XX			
XX			
CC	Clone A5 was prepd. from adult tick cDNA library. Clone 4U1 was		
CC	prepd. from the larval stage of B. microplus (Calliope strain).		
CC	AAQ82948 is a hybrid of sequences from clone 4U1 (nt 1-966 & 1747-2047)		
CC	and A5 (nts 967-1746). The translation of the tick carboxypeptidase		
CC	cDNA sequence is shown in AAR70013. All the native tick		
CC	carboxypeptidase sequences listed in Table 11 (see AAR70014-R70023)		
CC	are found in the translation (see FT). The predicted AA sequence		
CC	agrees with the peptide sequence for all peptides with 2 exceptions.		
CC	These differences are Asp for Glu14 in peptide T9126, and Asn for		
CC	Asp12 in peptide T9118. Tick carboxypeptidase has a mol. wt. 75172.		

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 16:56:13 ; Search time 20 Seconds
(without alignments)
3869.410 Million cell updates/sec

Title: US-09-978-385-2
Perfect score: 4291
Sequence: 1 MSSSWLLSLVAVTAQAOST.....ISKGNPQFQNTDVTQTSF 805

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4281	99.8	804	2	T14762	hypothetical prote
2	1344	31.3	732	1	S05238	peptidyl-di-peptida
3	1337	31.2	1306	1	A31759	peptidyl-di-peptida
4	1334	31.1	732	1	A35655	peptidyl-di-peptida
5	1334	31.1	1312	1	A34171	peptidyl-di-peptida
6	1312	30.6	1193	2	JC2489	peptidyl-di-peptida
7	1310	30.5	1313	1	JC2038	peptidyl-di-peptida
8	1283	29.9	737	1	A34402	peptidyl-di-peptida
9	1058	24.7	1309	1	S35484	peptidyl-di-peptida
10	1039	24.2	615	2	A57533	peptidyl-di-peptida
11	1030	24.0	630	2	JC5374	peptidyl-di-peptida
12	642.5	15.0	907	2	T15792	hypothetical prote
13	157	3.7	532	2	C83696	hypothetical prote
14	154	3.6	502	2	AF1310	probable thermosta
15	147	3.4	502	2	AE1682	peptide synthetase
16	139.5	3.3	987	2	AI2011	zinc metalloprotei
17	139	3.2	608	2	B82938	carboxypeptidase h
18	136	3.2	611	2	D82881	1,4-alpha-glucan b
19	135	3.1	501	2	S40048	G-utrophin - mouse
20	125	2.9	627	2	I48373	probable thermosta
21	124.5	2.9	987	2	E72561	conserved hypothet
22	124	2.9	538	2	E90270	lanthiotic epider
23	123.5	2.9	902	2	S23416	1,4-alpha-glucan b
24	123	2.8	642	2	E98000	oligoendopeptidase
25	122	2.8	607	2	AB3511	beta-galactosidase
26	121	2.8	607	2	T30574	probable SNF2 subf
27	121	2.8	1034	2	T30574	
28	121	2.8	1034	2	T30574	
29	120	2.8	1339	2	AB4663	

30	119.5	2.8	1283	2	S52500	SMH1 protein homol
31	119.5	2.8	3655	2	T38084	TRAP-like protein
32	119.5	2.8	4540	2	T30838	cytoplasmic dynein
33	118.5	2.8	1780	2	T17272	hypothetical prote
34	118	2.7	642	2	G95129	1,4-alpha-glucan b
35	117	2.7	1575	2	G82905	conserved hypothet
36	116	2.7	963	2	C90535	conserved hypothet
37	116	2.7	1642	2	T08880	NMDA receptor-bind
38	116	2.7	1939	2	D97316	probable S-layer p
39	115.5	2.7	611	2	A75573	probable oligendo
40	115.5	2.7	3433	1	S28381	utrophin - human
41	115	2.7	524	2	B82202	thermostable carbo
42	115	2.7	950	2	A71655	hypothetical prote
43	115	2.7	952	2	T50451	hypothetical colle
44	114.5	2.7	901	2	AB4653	phage infection pr
45	114.5	2.7	1225	1	B64234	hypothetical prote

ALIGNMENTS

RESULT 1

T14762
hypothetical protein DKFZp434A014.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #ext_change 20-Sep-1999

C:Accession: T14762

R:Wandut, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

A:Reference number: Z18181

A:Accession: T14762

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-804 <MAN>

A:Cross-references: EMBL:AL110224

A:Experimental source: adult testis; clone DKFZp434A014

C:Genetics:

A>Note: DKFZp434A014.1

Query Match
Best Local Similarity 99.8%; Score 4281; DB 2; Length 804;
Matches 803; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	2	SSSSWLLSLVAVTAQAOSTIEQATFLDKFHEHEDLFYSSLASMYNTNTEENVN	61
DB	1	SSSSWLLSLVAVTAQAOSTIEQATFLDKFHEHEDLFYSSLASMYNTNTEENVN	60
QY	62	MNNAGKMSAFLEOSTLAOMYPLQETQNLTKVLQALQALQNGSSVLSSEDSKRLNTLN	121
DB	61	MNNAGKMSAFLEOSTLAOMYPLQETQNLTKVLQALQALQNGSSVLSSEDSKRLNTLN	120
QY	122	TMSITVSTGKVCNPDPNPECLLEPGLNLTMANSLDYNRLMAMSENGKOLRPLYE	181
DB	121	TMSITVSTGKVCNPDPNPECLLEPGLNLTMANSLDYNRLMAMSENGKOLRPLYE	180
QY	182	EYVVLKEMARAHYEDYGDYRGDEYVNGVDYDSRQQLIEDYHTEETIKPLYEHLH	241
DB	181	EYVVLKEMARAHYEDYGDYRGDEYVNGVDYDSRQQLIEDYHTEETIKPLYEHLH	240
QY	242	AYVRAKLMNAYPSYISPIGCLPAHLGLDMGRFTWMLYSITVFGGKRPNTDVTDAVDDA	301
DB	241	AYVRAKLMNAYPSYISPIGCLPAHLGLDMGRFTWMLYSITVFGGKRPNTDVTDAVDDA	300
QY	302	WDAQRIFEAKRFVSGVGLPMTGFGFENSMILDPGVQKAAVCHPTAMDLGKDFRILMC	361
DB	301	WDAQRIFEAKRFVSGVGLPMTGFGFENSMILDPGVQKAAVCHPTAMDLGKDFRILMC	360
QY	362	TKVTMDDELTAHHEGHIOYDMAAOPFLRLRNGANGFEHANGVETISLSAATPKLKSI	421
DB	361	TKVTMDDELTAHHEGHIOYDMAAOPFLRLRNGANGFEHANGVETISLSAATPKLKSI	420
QY	422	GLSPDFQEDNETENFLKQALITVGTLPPTTYMLEKRWNVKGEIPKQMKKWEK	481
DB	421	GLSPDFQEDNETENFLKQALITVGTLPPTTYMLEKRWNVKGEIPKQMKKWEK	480

Db 421 GLLSPDFQEDNETEINFLKQALITVGLTPTFYMLEKRWVFKGELPKDQMKMKWEMK 480
 QY 482 REIVGVPEVPHDEYCDPASLFHVSNDYSFRTYTRLLYQFOEALCOAAKHEGPLHK 541
 Db 481 REIVGVPEVPHDEYCDPASLFHVSNDYSFRTYTRLLYQFOEALCOAAKHEGPLHK 540
 QY 542 CDISNSTEAGQKLFNMLRLGKSEPTLALENVGAKNNVPLNTEPEFTWIKDONK 601
 Db 541 CDISNSTEAGQKLFNMLRLGKSEPTLALENVGAKNNVPLNTEPEFTWIKDONK 600
 QY 602 SFVGSSTWSPYADOSIVRISLKSALGDAKAYENNDENMTLFRSSVAYAMKQYFLKVKNO 661
 Db 601 SFVGSSTWSPYADOSIVRISLKSALGDAKAYENNDENMTLFRSSVAYAMKQYFLKVKNO 660
 QY 662 MLFEEDVRYANLKPRISFNFVYAPKANSDIIPREVEKAIKMSRSRINDAFRLDND 721
 Db 661 MLFEEDVRYANLKPRISFNFVYAPKANSDIIPREVEKAIKMSRSRINDAFRLDND 720
 QY 722 LEFLGIQPTLGPPOPVSTWLYEGVYGVYVYIILFTGIRDRKKKKARSGENPY 781
 Db 721 LEFLGIQPTLGPPOPVSTWLYEGVYGVYVYIILFTGIRDRKKKKARSGENPY 780
 QY 782 ASIDISKGNNPGQNTDDVQTSF 805
 Db 781 ASIDISKGNNPGQNTDDVQTSF 804

RESULT 2

peptidyl-dipeptidase A (EC 3.4.15.1) precursor, testicular splice form - human
 N:Alternate names: angiotensin I-converting enzyme (ACE); CD143; dipeptidyl carboxypeptidase
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1991 #sequence
 C:Accession: S05238; #revision 02-Jul-1998 #text_change 18-Jun-1999
 R:Letton, A.L.; Soubrrier, F.; Allegrih, J.; Hubert, C.; Corvol, P.; Alhenc-gelas, F.
 FEBS Lett. 257, 99-104, 1989
 A:Title: The testicular transcript of the angiotensin I-converting enzyme encodes for the
 A:Reference number: S05238; MUID:89383720; PMID:2547653
 A:Molecule type: mRNA
 A:Residues: 1-732 <EHL>
 A:Cross-references: EMBL:X16295; NID:q28264; PIDN:CAA3362.1; PID:q28265
 R:Elhers, M.R.W.; Fox, E.A.; Stridom, D.J.; Riordan, J.F.
 Proc. Natl. Acad. Sci. U.S.A. 86, 7741-7745, 1989
 A:Title: Molecular cloning of human testicular angiotensin-converting enzyme: the testis
 A:Reference number: A33979; MUID:90046671; PMID:2554266
 A:Accession: A33979
 A:Molecule type: mRNA
 A:Residues: 1-732 <EHL>
 A:Cross-references: GB:M26657; NID:9338666; PIDN:AAA60611.1; PID:9338667
 A:Experimental source: clones R1.2 and T88
 C:Comment: neither the complete nucleic acid sequence nor the complete translation are shown
 C:Genetics: For the renal and pulmonary splice form, see PIR:A31759.
 A:Gene: GDB:DCP1; ACE
 A:Cross-references: GDB:119840; OMIM:106180
 A:Map position: 17q23-17q23
 C:Function:
 A:Description: catalyzes the hydrolysis of dipeptides from the carboxyl end of polypeptide
 C:Keywords: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; glycoprotein; metalloproteinase; peptidyl-dipeptide hydrolase
 F:1-21/Domain: signal sequence #status predicted <Sto>
 F:22-732/Product: peptidyl dipeptidase I #status predicted <MAT>
 F:666-702/Domain: transmembrane #status predicted <TRM>
 F:103,121,140,186,368,617,651/Binding site: carbohydrate (asn) (covalent) #status predicted
 F:414,418,434/Binding site: zinc, catalytic (His, His, Glu) #status predicted
 F:415/Active site: Glu #status predicted

Query Match 31.3%; Score 1344; DB 1; Length 732;
 Best Local Similarity 41.8%; Pred. No. 8,5e-85;
 Matches 259; Conservative 119; Mismatches 204; Indels 38; Gaps 10;

QY 15 TAAOS---TIEQAKTFLLDKFNHAEDELFGSSLASNMVNNITTEE-----NVQNM 62

Db 61 TAAQSPNLTDAEAEKSKVEEDRTSOVVMWEYEAANNVNTNTTETSKILLQKNQIA 120
 QY 63 NNAQKWSAFLEKQSTLAQMTPLQEIQLVYKLLQALQAGSSVLEEDSKRLNTLNT 122
 Db 121 NHT-----LKYGTAQKRFVDVNOQLNTTIKRIKKVQDLERAAALPAQLEELNLLD 172
 QY 123 MSTIYSGKVCQPDQPOQCLLEEGLEIANASLDYERLMAWESRSEVGKQRLPYER 182
 Db 173 METTYSVAIVCHPNS--SCDLEPDLITVMAITSKRYEDLLMAEGWMDKGRALILQFYPK 230
 QY 183 YVTLKNEKARANNHEDYGDYWRGDYEVNGVDYDYSRGQLIEDYEHFEIRKLYELHA 242
 Db 231 YVELINQARANGVYDAGDSWRSMTETPSLE-----QDLERLFOELQYILNHLA 280
 QY 243 YVRAKIMAY-PSYISPIGCLPAHLDDMGGRWTNLVSLTFPGKPNIDVTDAMVDA 301
 Db 281 YVRALIRHYGAQHTINLEGPIPAHLGNNMAQWMSNITDLYVFPSPASMDTTEAMLKG 340
 QY 302 WDAQRIKREAEKFFSVGLPMTQGFENSMITDPCNVOKAVCHPAPMDLGGK-DEPILM 360
 Db 341 WTPRMKREADDFTTSLGLLPVPEFMNKSMLKPTDGRVYCHASAMDEYKDFRIKQ 400
 QY 361 CTKYMDPFLAHHEMGIQYDAIYAAQPLRLRNGANEGFHEAVGEIMSLAATPKHLS 420
 Db 401 CTIYNLEDLVAHHENGHIQYFMQYKDLPAVLRGANGPFGHAIQDVALSVSTPKHLS 460
 QY 421 IGLSPDFQEDNETEINFLKQALITVGLTPTFYMLEKRWVFKGELPKDQMKMKWEM 480
 Db 461 LNTLSRGSGSD-BHDINFLMKMALDKIATIPSYLVQDKRWKVFSGSTKKEYNOEMWSL 519
 QY 481 KREIVGVPEVPHDEYCDPASLFHVSNDYSFRTYTRLLYQFOEALCOAAKHEGPLH 540
 Db 520 RLKYQSGCPVPTQGFDPFGAKFHTPSVPIRYFVSFTIOFQHEALCOAGHTGELH 579
 QY 541 CDISNSTEAGQKLFNMLRLGKSEPTLALENVGAKNNVPLNTEPEFTWIKDONK 600
 Db 580 KCDIYQSKBAGQRLATAMKLGFSRPMELMQLITQPMASASAMLSYFKPLDMLRTENE 639
 QY 601 --NSFYGM--STWSPYADOS 617
 Db 640 LHEGKLGMQYVMTWTPNSAKS 659

RESULT 3

peptidyl-dipeptidase A (EC 3.4.15.1) precursor, renal and pulmonary splice form - hum
 N:Alternate names: angiotensin I-converting enzyme (ACE); CD143; dipeptidyl carboxypeptidase
 C:Species: Homo sapiens (man)
 C>Date: 07-Jun-1990 #sequence
 C:Accession: A31759; P00004
 R:Soubrrier, F.; Alhenc-gelas, F.; Hubert, C.; Allegrih, J.; John, M.; Tregear, G.; C
 Proc. Natl. Acad. Sci. U.S.A. 85, 9386-9390, 1988
 A:Title: Two putative active centers in human angiotensin I-converting enzyme reveal
 A:Reference number: A31759; MUID:89071703; PMID:2849100
 A:Accession: A31759
 A:Molecule type: mRNA
 A:Residues: 1-1306 <SOU>
 A:Cross-references: GB:004144; NID:q178285; PIDN:AAA51684.1; PID:q178286
 A:Experimental source: kidney
 C:Note: parts of this sequence, including the amino end of the mature protein, were d
 R:Takeuchi, K.; Shimizu, T.; Ohishi, N.; Seyama, Y.; Takaku, F.; Yotsunoto, H.
 J. Biochem. 106, 4424-445, 1989
 A:Title: Purification of human lung angiotensin-converting enzyme by high-performance
 A:Reference number: P00004; MUID:90110025; PMID:2556109
 A:Accession: P00004
 A:Molecule type: protein
 A:Residues: 'XX', 32-34, 'E', 36-37, 'X', 39-41, 'R', 43-46 <YAK>
 A:Experimental source: lung
 C:Comment: This splice form is found in many tissues, in particular kidney and lung v
 C:Genetics:
 A:Gene: GDB:DCP1; ACE
 A:Cross-references: GDB:119840; OMIM:106180

A:Cross-references: GB:U03734; NID:9437289; PIDN:AAA8211.1; PID:9437290
 A>Note: the authors translated the codon ACC for residue 159 as Tyr
 C:Comment: This enzyme is a zinc-containing dicarboxy peptidase that cleaves angiotensin I
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; peptidyl-dipeptide hydrolase; transmembrane protein; 24
 F:393-400,990-998/Region: catalytic status predicted
 F:1284-1284/Domain: transmembrane status predicted <TM>

Query Match 30.5%; Score 1310; DB 1; Length 1313;
 Best Local Similarity 42.0%; Pred. No. 4,7e-82;
 Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps 7;

20 TEEQAKFLDKFNHEADLFYQSSIASMWNNTNTEENONMANNAGDGSAFLKEOSTL 79
 DB TDEKANKFVEYDRTAVLNVEAEANMHNNTITGSKILLQKNVEYNHLKTYGM 709
 QY 80 AOMVPLQEIQLVYKQLQALQNGSSVLSDESKRLNTLNTSTYSTGKVCNPDPQ 139
 DB 710 AKTEVSNFQNSTIKRIKKYQNDRAVLPPNELEEVNQILLDMETYSVANVCYNG-- 767
 QY 140 ECLLEPELNEIMNSLDYNERIAMESRSEVGKQLRPLKEEVYVLKNEMARAHYEDY 199
 DB 768 TOLSEPLNIMATSRKYEELLMWYKSRKVGKRALPPKPYVESNKLAKNGYSDA 827
 QY 200 GDYWGDEYVNGVDGYDSRGLIEDVEHTEIEIKPLYEHLHAYVRAKIMAYPS-YISP 258
 DB 828 GDSMWSYSESDLE-----QDLKLYQELQPLVNLHAYVRSRLHHYGESEYINL 877
 QY 259 IGLCPAHLGLDMGFRFTNLVSLVPGOKPNIDVTAMVDQADQRIKFEAEKFEVSV 318
 DB 878 DGPPIPAHLGNMMAQTWSNIYDLVAPFPASPIDATEAMIKQWTPRIKFEADNFTSL 937
 QY 319 GLPNTOGFENSMGLTDGYNOKAVCHPTAMDLGK-DERILMCTKYMDPEFLAHNMG 377
 DB 938 GLPPPEPEFNKSMLEKPTDGEVYCHASAMDFNGKQFRIKQCTSVNNEELVLAHBMG 997
 QY 378 HLYDMAVAAAPFLLRNGANEHGEVGEIMSLSAATPKHKSIGLSLSPDEQENETEN 437
 DB 998 HLYQWQKDLPTVREGANFGHEALIGDVALSYTPKHLHSLNLSS- GSGEHNIN 1056
 QY 438 FLKQALITVGLPTFTMLEKRMVYFKGEPKIDQMKMKMKREIYGVVPPHDETY 497
 DB 1057 FLKMAALDKIAFIPSYLIDDMRMVFGSITKENYNOEMWSLRKLYGGLPPVPRSGD 1116
 QY 498 CDPALFVSNDSYFIRYRTLYQFOFQALCOAKHEGLKCDISNTEAGKLENM 557
 DB 1117 FDPGSKFHPANVPYIRFISITIOFQHEALCRAAGHTGPKYKCDIYQSKEDAGLLADA 1176
 QY 558 IRLGSEFWTLALENVGAKNNVAPLLNTEPELFTWLKDQK--NSFVGN-STDWSP 612
 DB 1177 MRLGSKQWPEAMKITIGOPNMSASAINNFKPLTEMLVTENRRRGELTGWPEYTWTP 1234
 RESULT 8
 A34402
 N:Altemate names: angiotensin I-converting enzyme; dipeptidyl carboxypeptidase I; pepti
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A34402; A60724; A36232; C18700
 J:Kumar, R.S.; Kusari, J.; Roy, S.N.; Softer, R.L.; Sen, G.C.
 J: Biol. Chem. 264, 16754-16758, 1989
 A:Title: Structure of testicular angiotensin-converting enzyme. A segmental mosaic isozy
 A:Reference number: A34402; MUID:89380303; PMID:2550457
 A:Accession: A34402
 A:Molecule type: mRNA
 A:Residues: 1-737 <KDD>
 A:Cross-references: GB:U05041; NID:9164744; PIDN:AAA31153.1; PID:9164745
 J:Sen, G.C.; Thekkumkara, T.J.; Kumar, R.S.
 J: Cardiovasc. Pharmacol. 16(Suppl.4), S14-S18, 1990
 A:Title: Angiotensin-converting enzyme: structural relationship of the testicular and the
 A:Reference number: A60724; MUID:9115372; PMID:1705622

A:Accession: A60724
 A>Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 73-173 <SEN>
 A>Note: Identical sequences were obtained for mRNAs from lung and testes
 R:Chen, Y.N.P.; Riordan, J.F.
 Biochemistry 29, 10493-10498, 1990
 A:Title: Identification of essential tyrosine and lysine residues in angiotensin con
 A:Reference number: A36232; MUID:91104959; PMID:2176870
 A:Accession: A36232
 A:Molecule type: protein
 A:Residues: 154-160,236-242 <CHE>
 R:Iwata, K.; Lai, C.Y.; El-Dorry, H.A.; Softer, R.L.
 Biochem. Biophys. Res. Commun. 107, 1097-1103, 1982
 A:Title: The NH2- and COOH-terminal sequences of the angiotensin-converting enzyme iso
 A:Reference number: A90107; MUID:83048249; PMID:6291514
 A:Accession: C18700
 A:Molecule type: protein
 A:Residues: 33-35,'SN','SS','FAEL','737 <IWA>
 A>Note: several of the amino acids in reported are tentative
 C:Comment: The pulmonary and testicular isoforms of this enzyme differ substantially
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; peptidyl-dipeptide hydrolase; testis; transmembrane

Query Match 29.9%; Score 1283; DB 1; Length 737;
 Best Local Similarity 40.8%; Pred. No. 1.4e-80;
 Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;

20 TIEQAKFLDKFNHEADLFYQSSIASMWNNTNTEE-----NVQNMNN--AGDKW 69
 DB 75 TDEKANKFVEYDRTAVLNVEAEANMHNNTITGSKILLQKNVIAHNTLTYGHW 134
 QY 70 SAPLKESTLAQWPLQEIQLVYKQLQALQNGSSVLSDESKRLNTLNTSTYST 129
 DB 135 -----ARRDVSNFQNSTIKRIKKYQNDRAVLPPNELEEVNQILLDMETYSV 184
 QY 130 GKYCNPNPQECLELLEPELNEIMNSLDYNERIAMESRSEVGKQLRPLKEEVYVLKNE 189
 DB 185 ANNCRVG--SCQLDEPLDNLNMAATSKRYDELMLVMTSMWDXKGRALILPFPYVFTNK 242
 QY 190 MARAHYEDYGDYWRGDEYVNGVDGYDSRGLIEDVEHTEIEIKPLYEHLHAYVRAKIM 249
 DB 243 AARLNGYVDADGDSRSRYETPTLE-----QDLERLFOELQPLVNLHAYVRAH 292
 QY 250 NAY-PSYISIGCPAHLGLDMGFRFTNLVSLVPGOKPNIDVTAMVDQADQRIK 308
 DB 293 RHGAQHINLEGPPIPAHLGNMMAQTWSNIYDLVAPFPASPIDATEAMIKQWTPRIK 352
 QY 309 KEAEKFEVSVGLPMTQGFENSMGLTDGYNOKAVCHPTAMDLGK-DERILMCTKYMD 367
 DB 353 EADKFTISGLPLPPPEPEFNKSMLEKPTDGEVYCHASAMDFNGKQFRIKQCTSVNNE 412
 QY 368 DELTAHHEMGIQDMAVAAAPFLLRNGANEHGEVGEIMSLSAATPKHKSIGLSLSD 427
 DB 413 DLVYVHEMGIQDFMYKDLPAALREGANFGHEALIGDVALSYTPKHLHSLNLSS- 472
 QY 428 FQEDNEINFLKQALITVGLPTFTMLEKRMVYFKGEPKIDQMKMKMKREIYGVVPPH 487
 DB 473 -GGYEHDINFLKMAALDKIAFIPSYLIDDMRMVFGSITKENYNOEMWSLRKLYGGL 531
 QY 488 VEPYHDETYCDPASPALFVSNDSYFIRYRTLYQFOFQALCOAKHEGLKCDISNS 547
 DB 532 CPVAPRSGQGFDPGAKNHIPSPYIRFISITIOFQHEALCRAAGHTGPKYKCDIYQ 591
 QY 548 TEAGQKLFNMLRLGSEFWTLALENVGAKNNVAPLLNTEPELFTWLKDQK--KNSFVG 605
 DB 592 KEAGKRLADAMKLGYSKPEAMKITIGOPNMSASAINNFKPLTEMLVTENRRRGELT 651
 QY 606 W-STDWSPYADQS 617
 DB 652 WPOYTWTPNSARS 664

[illegible]

RESULT 11

peptidyl-dipeptidase A (EC 3.4.15.1) 67k precursor - fruit fly (*Drosophila melanogaster*)
 N:Alternate names: angiotensin converting enzyme
 C:Species: *Drosophila melanogaster*
 C:Date: 08-Feb-1996 #sequenceRevision 08-Feb-1996 #textChange 26-Feb-1998
 C:Accession: A57533
 R:Connell, M.J.; Williams, T.A.; Lamango, N.S.; Coates, D.; Corvol, P.; Soubrier, F.; HC
 J. Biol. Chem. 270, 13613-13619, 1995
 A>Title: Cloning and expression of an evolutionary conserved single-domain angiotensin c
 A:Reference number: A57533; MUID:95293950; PMID:7775412
 A:Accession: A57533
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-615 <COR>
 A:Cross-references: GB:U25344
 C:Genetics:
 A:Gene: FlyBase:Ance
 A:Cross-references: FlyBase:FBgn0012037
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: peptidyl-dipeptide hydrolase

Query Match	24.2%;	Score 1039;	DE 2;	Length 615;
Best Local Similarity	35.8%;	Pred. No. 7.8e-64;		
Matches 219;	Conservative 120;	Mismatches 251;	Indels 22;	Gaps 9;

QY 8 LLSLVAAYAAOSTIEEQAKTEFDKRNHHAEDLFQSSLASNNYNTTTEENVOOMNAGD 67

8 LLSLVAAYAAOSTIEEQAKTEFDKRNHHAEDLFQSSLASNNYNTTTEENVOOMNAGD 67

Db 8 LLAATLAVGQALKEEIOAKEYDENLNENKLAKRTNVETFAAWAYSNTDENKKKNISA 67

QY 68 KMSAPLKEBSTLAQMYPILOEIONLTVKLTOLALOONSYSVEGSKRKLNTLMTWTY 127

68 KMSAPLKEBSTLAQMYPILOEIONLTVKLTOLALOONSYSVEGSKRKLNTLMTWTY 127

Db 68 ELAKFEKVEASDTTFQWRKSIOSEDLKQFALTKLGYALPEDDYALDLTSLAMSNF 127

68 ELAKFEKVEASDTTFQWRKSIOSEDLKQFALTKLGYALPEDDYALDLTSLAMSNF 127

[illegible]

RESULT 12

angiotensin-converting enzyme-related protein - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C:Date: 28-May-1997 #sequence_revision 18-Jul-1997 #text_change 21-Jul-2000
C:Accession: J05374
R:Taylor, C.A.M.; Coates, D.; Shitras, A.D.
Gene 181, 191-197, 1996
A:Title: The Acer gene of *Drosophila* codes for an angiotensin-converting enzyme homo.
A:Reference number: J05374; MUID:97128790; PMID:8973330
A:Accession: J05374
A:Molecule type: mRNA
A:Residues: 1-630 <RAY>
A:Cross-References: EMBL:X96913; NID:g1405881; PIDD:CA65652.1; PID:g1405882
C:Genetics:
A:Gene: Acer
C:Superfamily: mammalian peptidyl-dipeptidase A

Query Match	24.0%	Score 1030;	DB 2;	Length 630;
Best Local Similarity	35.6%;	Pred. No. 3.4e-63;		
Matches 219; Conservative 113; Mismatches 260;		Indels 24;	Gaps 10;	

Query Match	24.0%;	Score 1030;	DB 2;	Length 630;
Best Local Similarity	35.6%;	Pred. No. 3.4e-63;		
Matches 219;	Conservative 113;	Mismatches 260;	Indels 24;	Gaps 10;

```

Oy  6 WLLSLVAVTAQSTIEQAKFLDKFHEHEEDLFYOSLASVMYNNINIEVQJNNNA 65
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db  16 WLFPHGISMGNSCSASVLE-ARREFLEMBQLRRRHFHEFLSGYNNINYEANQAMIEV 74
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Oy  66 GKRWSAFLEKOSTLAQMYPLOEIQMLTVKLOLQALOONGSSVLESDKSKRLNTLNTMST 125
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db  75 YARNELMNLRLAQOIKSSDPYVOSEPADIRROAEHLSTKGASALNADYDYLALQNAISSMOT 134
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Oy  126 IYSTGVCNCPADPOEC-LLEFGLMEIWMANSLDYNERLMAESWRSEVQKOLPBYEEV 184
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db  135 NATATAYVCYTTRKSDCSLLEPHIOERLSHRDPAELAWYREKHNDKSGTRPMQNEFEV 194
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Oy  185 VLKNEMARANHEDYGDYWRGDEYVNGVDYDYSRGQLLEDVEHTFEELIKPLVHLHAY 244
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

```

Db 195 RLEKASQNLGHRSYADYVQFYE-----DPEFER-----QLDAIFKQILLPLRYQLHGVY 244
 Qy 245 RAKLMNAY--PSYISPIGCLPAHLIGDMGKFWNLISLYVFGCKNIDVTDAVDQAMD 303
 Db 245 RFRROHIGDVPABESNIPISLIGNMWGSNNELLDLPYEPKFPVYKAKEMEKQY 304
 Qy 304 AORIFKEAEKFPFVSVGLPNTQGFVNSMLTDPGNOKAVCHPTADLCK--GDFRILMCT 362
 Db 305 VQRLFELGQGFQSLGRALPSPFWMLSVLTRDD--KQVYCHASAMDEFQDSQVRLKCT 363
 Qy 363 KATMDPLTAHMHMGHIQYDMAAOPFLLRNGANGFHEAVGEINLSAATPKHLKSTG 422
 Db 364 EVDSHFYVYVHHELGHIQYLYEQOPAVYRGAPNGFHEAVGVYALSVMSAKHLKALG 423
 Qy 423 LLSPEDEDETEINFLKQALITVGLTPTYLLEKRWVWVKGELPKQWKKWEMK 482
 Db 424 LIE--NGRLDEKSKINOLEFQALSKITVLPFGAVDVKYRAVRNEDESQWNCGEQWMS 482
 Qy 483 EIVGVVEVPVPHDETCYDPAFLFVNSDYFIRYRTLYQFOFQALCOAKHEGP---- 538
 Db 483 EFGVPEPVPYFTEKEDPDPAKYHIDADVEYLRFAHIGFQFHKYLCKKAKQYAPNNSR 542
 Qy 539 --LHKDINSSTPAGOKLENNMLRLKSESWTALENVVGAKKMANNRPLINFEPLFTWLK 596
 Db 543 LITDNCDDIFGSKAAGRSLSQFLSKGNSRMKKEVLEETGETEMDPAALIEFPEPLYOWLK 602
 Qy 597 DONKNSFVGWSTWSP 612
 Db 603 QE--NSRLGVPLGMP 616

RESULT 13

115792
 Hypothetical protein C42D8.5 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
 R:Hallsworth, K.
 Submitted to the EMBL Data Library, April 1996
 A:Description: The sequence of C. elegans cosmid C42D8.
 A:Reference number: 218405
 A:Accession: U15792
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-907 <HAL>
 A:Cross-references: EMBL:U159666; NID:G1293844; PID:G1293847; PIDN:AAA98719.1; GSPDB:GN00
 A:Experimental source: strain Bristol N2; clone C42D8
 A:Genetics:
 A:Gene: CESP:C42D8.5
 A:Map position: X
 A:Introns: 140/3; 170/3; 194/3; 300/2; 467/3; 551/2; 600/2; 697/3; 774/2; 851/3

Query Match 15.0%; Score 642.5; DB 2; Length 907;
 Best Local Similarity 27.0%; Pred. No. 3.6e-36;
 Matches 172; Conservative 136; Mismatches 289; Indels 41; Gaps 19;

Qy 2 SSSSLLLSIYAVTAAGSTIEQAKTFLDKFNHEADLFTQSSLASWNTNTITEENVQ 61
 Db 160 SSNTKTDNLDPGSIK--BEKLRSWLAGYEAKYLRVALSGRWYFNDASPSIKLA 217
 Qy 62 MNNAGDKWASFLKEQSTLAQWYPLQETQNTLVKTLQALQNGSSVLSSEKSKLNTILN 121
 Db 218 IDEAVNLTMFVRSQMAKOFDASVYDEKVMQGLGVSEGGASALAPSFADYSAQA 277
 Qy 122 TMSITISGKVCNPDNPOECILLPEGLNETMANSLDYNERLWAMESRSEVQKRLPYE 181
 Db 278 ALNRDSKSTICDKDVPKALQKIDMSIFRNEKSDASRLQHLWVSVYTAIAKS--KSYN 336
 Qy 182 EYVVLKNEAMARANYEDYGDYWRGDYEVNG--VDCYDYSKQGLIDVYHTTEIKPLTEHL 240
 Db 337 NIITISSEGAFLNFGANGAMRSAPFMSVSKHAEF--DLNKQIDKISTIOFPYOL 393
 Qy 241 HAYVRAKLMNAY--PSYISPIGCLPAHLIGDMGKFWNLISLYVFGCKNIDVTDAV 298

Db 394 HAYMRQLAGTISNPFVSKDGPPIPAHLFGSLDGGDMAHAEQKPFEEES--ETPEAML 451
 Qy 299 D-----QAMDQORIFKEAEKFPFVSVGLPMMTQGFVNSMLTDPGNVQAVCHP--TAMPL-G 352
 Db 452 SAFNTQNTYTKKMFATATIRKSAEPPLPSPYMTSSIFARWVS--KMDICHPAALDMRA 510
 Qy 353 KGFDFILMCTRYTMDPLTAHMHMGHIQYDMAAOPFLLRNGANGFHEAVGEINLSA 412
 Db 511 PNDFFVKACAOQGEDEFQASLSLVQYLYQLYKDSQLLFFREQASPYITDAIARFHL 570
 Qy 413 ATPKHLKSTIGLSPQEDNTE--INFLKQALITVGLTPTYLLEKRWVWVKGELPKD 471
 Db 571 TNPFLYQKLVPSSEHLDIKDSVINIKLYKESLSPFKLPTTIADWMMRYELFGQVAPKN 630
 Qy 472 QMKNWEMKREIIVGVPEVPHDETCYDPAFLF--VSNDSYFIRYTRTL---YQFQF 525
 Db 631 KLNDRWMEIRMKYEGVRSRPOYNTSND--ALHNSVQVHS--PATRLLISLYLKFQI 685
 Qy 526 QEALCOAA--KHGAPLHKDISNSTEAGOKLENNMLRLKSESWTALENVVGAKKMANNR 582
 Db 686 LKALCQRELFWLSBG----CILSDYT--EKLRETMKLGSSITWKLALEMISKGELDAQ 739
 Qy 583 PLNVFPEPLFTWLKDONK--NSFYGWSGMSDVPYADQSI 618
 Db 740 PLLEYEPFLINWLNKNTNEIDQVYVGVGDEGTPVEEI 777

RESULT 14

C83696
 Hypothetical protein BH0371 [imported] - *Bacillus halodurans* (strain C-125)
 C:Species: *Bacillus halodurans*
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a
 A:Reference number: A83650; MIMD:20512582; PMID:11058132
 A:Accession: C83696
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-532 <STO>
 A:Cross-references: GB:AP001508; GB:BA000004; NID:G10172890; PIDN:BAH04090.1; GSPDB:G
 A:Experimental source: strain C-125
 A:Genetics:
 A:Gene: BH0371

Query Match 3.7%; Score 157; DB 2; Length 532;
 Best Local Similarity 21.1%; Pred. No. 0.0058;
 Matches 118; Conservative 83; Mismatches 213; Indels 144; Gaps 29;

Qy 22 EDAQKTFDKFNHEADLFTQSSLASWNTNTITE--NVQNNMAGDKWASFLKEOS--- 77
 Db 3 EODIERFLSQNRVBDLPQVLLNMWVATGDEQMSQHSLSSEYWAHSDESSEFOK 62
 Qy 78 -----TLAQMYPLOEIQNTLVKTLQALQNGSSVLSSEKSKRLNTILNTMSTLY 127
 Db 63 VTRFRKIDSLPLKQRQQLDDHDKMIKNQFE--EGRQOILDE--KTSHYFTFPQOY 118
 Qy 128 STGKVCNPDNPOECILLPEGLNETMANSLDYNERLWAMESRSEVQKRLPYEYVYLK 187
 Db 119 NSGRVSNNE-----LIDLIRYDDHRRKQAWPA--SKEVGKRTKEDLLQILRR 166
 Qy 188 NEMARANYEDYGDYWRGDYEVNGVYSGQGLIEDVEH---FEELKRYELHLNAVY 244
 Db 167 NEVARNLGFTF-----YMSRATQELDEQTFAMFETIKSSDDAFRMI 211
 Qy 245 -----RAKLMNAYPSYISPIGCLPAHLIGDMGKFWNLISLYVFGCK--PNIDVTDA 296
 Db 212 KDEIDERAIVKIKKDDLRF-----WDYVDPEFGQAPSTEIEND- 250
 Qy 297 MVDQAMDQORIFKEAEKFPFVSVGLPMMTQGFVNSMLTDPGNVQ--AVCHPTAMDLGKD 355

Db 251 -FDSEFKDQDLEQVVSQTFQAMLP--IDDILKRSULYPRKNKNPFQFC--TDMQ-RKQD 304

Qy 356 FRIIMCTKVTMDLFLAHMHMGH-IOYDMAAOPFLLRNGANEFGHEAVGEINLSAAT 414

Db 305 IRYILINDQSMYWTALHHEFGNAVYKFDLSLPFLRL-----FH-----SHTLTT 351

Qy 415 PKHLKSIGLS--PDQGE-----DNET-----EINFLKQALTIYGLTFYMLEKRW 461

Db 352 EASALFGRMTKMAEWYERFLGIDRETCERIGRMEMKLRQM-VYST-----RW 400

Qy 462 MV----FKG---ELPKQOMKKMKWEMKREITVGVEYPHETCDPASLFHVS-----N 508

Db 401 MLTFEFKSLYEDPDODINALMKLVKELQYMAP--PEDTGSPPDMAAKMHFSLAPVYQ 458

Qy 509 DY-----SEIRYRT 519

Db 459 DYLLGEMASOLHRYTKT 476

RESULT 15

AF1310

Probable thermostable carboxypeptidases homolog lmo186 [imported] - Listeria monocytogenes

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AF1310

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Me

ok, C.; Schluter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund, A.

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MID:21537279; PMID:11679669

A:Accession: AF1310

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-502 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAC99964.1; PID:g16411339; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo186

C:Superfamily: Thermus aquaticus carboxypeptidase Tag

Query Match 3.6%; Score 154; DB 2; Length 502;

Best Local Similarity 20.1%; Pred. No. 0.0086;

Matches 128; Conservative 101; Mismatches 215; Indels 192; Gaps 35;

Qy 20 TIEBOAKTFLDKFNHEADLEFYQSSLASWNYNT-----NITEENVQNMNA 65

Db 4 TLEEFLLAYIKKWEALEBAL-----ALYWDLRTGAPAKMEGRSDVIGVLSSEIFNMQTS 59

Qy 66 GDKWSAFLKEOSTLAQMYPLQEIQNLTVKQLQALQNGSSVYSEKSKRLT-----N 117

Db 60 -EEMAAFLAGLN-----QDKENLS-ETTRKTLLE--SQKTYDLNKKIKSEYAEY 106

Qy 118 TILMTSTYSTGKVCNDPQECLELLEPGINEMANSIDYNERLMAWESWRSEVGKQLR 177

Db 107 KLVAAQEAETWTTARQN-----DRAAFEPPLTKL----- 136

Qy 178 PLYEYVVLKNEMARAHYEDYGDYWGDEYVGVG--DYSKQLIEDVEHTFEELIKP 235

Db 137 -----EMKR-----KFEVEY--GYENKYDILLDYEYGVVSYLDSVFEKVR- 177

Qy 236 LYEHLLHAVYRAKLMNAVSYISPIGCLPAHLGLGMMGRFTNLXSLVPPGQKRNIDYTD 295

Db 178 -DGTMA-IREKTIENE-----GVKPDATILN 200

Qy 296 AMVDQAMDQRIFFKAERFEVSVGLPNNTOGFWENSMILDPGNVQAKVCHPTANDLGKGD 355

Db 201 TKISA-----KQKESIRI-LNKKGFDF-----EAGRIDERY-HPFATGLMTGD 243

Qy 356 FRIIMCTKVTMDLFLTA-----HHEMGHIQY-----DMAVAADPFLLRNGANGFHEAVG-- 405

Db 244 VRI--TTRYNENDFKMAVFGTTHSGHAIYEQNDALVGP--LANGASGIESOSLF 299

Qy 406 -EIM--SLSAAPPKHLKSIGLSLSPDFO-----ED-----NETEINFLKQALTIYGLP- 451

Db 300 YELIIGSSIAFKSVYADPQATKPAFOVKLEDEYRAVNVSESSLINLEADTL--TYPL 357

Qy 452 ---FTYMLEKRWVYFKGEIPKQOMKKMKWEMKREITVGVEYPHETCDPASLFHVS 508

Db 358 HIMIRELEK---ALINGELEKDLPKAMGDYEEYLG1---RPNDNGVLODIHWAG 411

Qy 509 DYSFIRYRTLL-YQFQFQALCQAAKHGEPRLHKCDISNSTEAGQKLF--NMLRLKSE 564

Db 412 DGYEPSPALGLMYAAQFENQW---QKEIPIDAILIISDDYSELKTLTHVHKFGTK 467

Qy 565 PWTALENVYAKNMNVREPLNVEPPLFTWLKQNK 600

Db 468 KPELITLDTTG--EGINPTYLDDLKRRAYVYQFVK 502

RESULT 16

AE1682

Probable thermostable carboxypeptidases homolog lin199 [imported] - Listeria innocua

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AE1682

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Me

ok, C.; Schluter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MID:21537279; PMID:11679669

A:Accession: AE1682

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-502 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97229.1; PID:g16414500; GSPDB:GN00178

A:Experimental source: strain C11p11262

C:Genetics:

A:Gene: lin199

C:Superfamily: Thermus aquaticus carboxypeptidase Tag

Query Match 3.4%; Score 147; DB 2; Length 502;

Best Local Similarity 19.4%; Pred. No. 0.026;

Matches 124; Conservative 95; Mismatches 223; Indels 196; Gaps 33;

Qy 20 TIEBOAKTFLDKFNHEADLEFYQSSLASWNYNT-----NITEENVQNMNA 65

Db 4 TLEEFLLAYIKKWEALEBAL-----ALYWDLRTGAPAKMEGRSDVIGVLSSEIFNMQTS 59

Qy 66 GDKWSAFL-----KQOSTLAQMYPLQEIQNLTVKQLQALQNGSSVYSEKSKRLT 119

Db 60 -EEMAAFLAGLNIDKEN-----LSEITRKTLLE-ESKTYDLNKKIKSEYAEYTKLV 109

Qy 120 LNTMTSTYSTGKVCNDPQECLELLEPGINEMANSIDYNERLMAWESWRSEVGKQLR 179

Db 110 AQA-ETWTTARQN-----DRAAFEPPLTKL----- 136

Qy 180 YEEYVVLKNEMARAHYEDYGDYWGDEYVGVG--DYSKQLIEDVEHTFEELIKP 237

Db 137 -----EMKR-----KFEVEY--GYENKYDILLDYEYGVVSYLDSVFEKVR-- 177

Qy 238 EHLHAVYRAKLMNAVSYISPIGCLPAHLGLGMMGRFTNLXSLVPPGQKRNIDYTD 297

Db 178 DGTMA-IREKTIENE-----GVKPDATILN 202

Qy 298 VQDQAMDQRIFFKAERFEVSVGLPNNTOGFWENSMILDPGNVQAKVCHPTANDLGKGR 357

Db 203 ISBA-----KQKESIRI-LNKKGFDF-----EAGRIDERY-HPFATGLMTGDVR 245

Qy 358 IIMCTKVTMDLFLTA-----HHEMGHIQY-----DMAVAADPFLLRNGANGFHEA----- 403

Db 246 I--TTRYENNDKMAVEGTIHGGAHYEDNDALVGP--LANGASMGHESQSLFYE 301
 Oy 404 -----VGEIMLSAATPRKHLKSIGLLSPDEQNETELNFKALITVGL 450
 Db 302 IITGSSLAWSKYNADFOAITKPADHYK----LEDFYKAVNISSSLRIPADTL--TY 355
 Oy 451 P----FTYMLERKRWVVFGEIPKDOMKKMKMKREIYGVPEVPHDEYCDPASFHFV 506
 Db 356 PLHIMIRYELEK---ALINGELEVKDLPKAMGDKEYEYIGI---RPNDNTNGVLODIIHA 409
 Oy 507 SUNDYRIRYRTFL-YQFOFOALCOAAHESPLKCDISNSTEGQKLF---NNLRIGK 562
 Db 410 GGDGFGFYSYALGLMAYAAQFYHOM---OKELIPNDIATIASDNYTELKWTIKVHTEGK 465
 Oy 563 SEPWTLALENVYGAKNMNVRLNTEPEPLFTWLKDONK 600
 Db 466 TKKPLEILTDITG-EGINFTYLLDLLEKRYAVYQFNK 502

RESULT 17

A12011

peptide synthetase [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: A12011

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriduchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: A12011

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-987 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA078013.1; PID:917135467; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

Query Match 3.3%; Score 139.5; DB 2; Length 987;
 Best Local Similarity 19.6%; Pred. No. 0.24; Mismatches 208; Indels 191; Gaps 26;
 Matches 117; Soches 80;

Oy 52 TITEENYQNNMAGDKMSAFLEQSTLAQMPLOEIONLVKLOLQALQNGSSVISED 111
 Db 106 TLITHEQVSOQLVACVQALPLQYVFLDEGPPLGEIQSLT-----QITAFWQS-LSDD 159
 Oy 112 KSKRLNTLNTMSTIYSGKVCNPN-----POCLLEPGLNETMANSLD 157
 Db 160 ALELCNSPDLLMVLITLTSGSTGRKGVMLNHRGYMRLTMQNTFSLQPGDRAQRTSFC 219
 Oy 158 YNERLAWESMRSEYKQLRPLYEYVVLKNDMARANHYEDYGDWRGDYVNGVDGDY 217
 Db 220 FDISWVE--IFPTIMSGAITCPQKRVVLANPEFAR-----WIDETQINWVHVPVS 268
 Oy 218 SRGQLEDEVEH---TFEEIKPLYEHLNAYRAKLINAYPSYISPLGCLPAHLILGDMGRF 274
 Db 269 LFGEFISALENETWSPQLRMLMFSGEA-----LPMSEFI-----QR 304
 Oy 275 W-----TNLYSLTVPGCKNDIVTDAMVDQANDAO-----RIKFAEKFEVSV- 318
 Db 305 WIDRHGLTGLANLYGPT-----EASIDIVCHLITERDEBLTQIPIGKAIDWVYKVL 359
 Oy 319 ---GLPNNMVG---FWNSM-----LFDPGNVQKAVCHPAMVLGSD--FRILMCTKV 364
 Db 360 DGMGPVPGNNGELMLGCVOLAGYLDPEKTPAQAFCPNPFTPI--PDYIYRTGDLVKE 418
 Oy 365 TMDDELTAHENGAIQYDMAYAAQPFLLRNGANGFGEHVAEYISLSAATPKHLKSGIL 424
 Db 419 LPDGTIEYH---GRIDHQVKI-----RGFRLEGEIESVLTTHP--DVREAAAL 462
 Oy 425 SPDFQEDNETELNFKALITVGLTPTTYMLEKRWVVFGEIPKDOMKKMKMKREI 484

Db 463 AVDIGEGOKRLV-----ACISGKTKIKRPFKEYLEOK--- 494
 Oy 485 VGVVEPVPHDEYCDPASLFHVSNDYSFIRYTYRTLYQFOFOALCOAAHESPLKCDI 544
 Db 495 -----LPH---YMLP-----ORFLMLDSL--PKHNKGLKLBKAL 523
 Oy 545 -----SNSTEAGOKLFMNLRLGKSEPTTLAENVYGAKNMNVRLNTEPEPLFTWL 595
 Db 524 VTQLTSDSPSSPLRLPLPLGPAQRW-----LVKYEFPYQWL 562

RESULT 18

B82938

zinc metalloprotease oligoendopeptidase F U0065 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000

C:Accession: B82938

R:Glass, J.I.; Leftkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to Genbank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a

A:Reference number: A82870

A:Accession: B82938

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-608 <GLA>

A:Cross-references: GB:AE002106; GB:AF222894; NID:g6899011; PIDN:AAF30470.1; GSPDB:GN

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: pepF-1; U0065

A:Genetic code: SGC3

Query Match 3.2%; Score 139; DB 2; Length 608;
 Best Local Similarity 19.1%; Pred. No. 0.13; Mismatches 226; Indels 240; Gaps 32;
 Matches 132; Conservative 93;

Oy 22 EQQAK-----TFLDKENHEAE-----DLFYQSSLAWSYNTNITEENYQNNM 63
 Db 28 EKKAKYIKAFPTFLDSKQNFQWQILIEFTIVANRF--NYVSNLNNTNVVD----- 78
 Oy 64 MAGDKMSAFLEQSTLAQMPLO-----ELQNLTVKLOLQALQNGSSV- 107
 Db 79 ---PKMSW---SOKLSAFYELETALSNVSVLANEAKIKETVLSQNLVYTRKNEYIF 133
 Oy 108 -----LSEPKSKRLNTLNT---MSTYRST-----GKVCNPNDEQCL 143
 Db 134 RYQPHILANQSKLPSLILADGCFSTIRTYNNDMKSDAIDAGKHPKIKNEAEFV 193
 Oy 144 LEPGLNETMANSLDYNERLAWESMRSEYKQLRPLYEYVVLKNDMARANHYEDYGDY 203
 Db 194 HLKSKDRVLKSA--YLSMRYAYDSRESITKM---LYYVLSL--NQAAKANNDDV--IA 246
 Oy 204 RGYEVNAGVGYDYSKQLEDEVEHFFEEKPLYEHLNAYRAKLINAYPSYISPLGCLP 263
 Db 247 KAAFD---DVIDKSLITLLYDQVKLYKDTNEDYKRVRTYLLKILKV--SKIEP----- 295
 Oy 264 AHLIGDMGRGFNTLSLVYFGCKNDIVTDAMVDQANDAOIRFEAEKFEVSVGLPMV 323
 Db 296 -----WNGPLP---ISKTIIDPIEEK-OMALDISLSIGDE---YVS-----NI 333
 Oy 324 TQGFWENSMLTDPGNVQKAVCHPAMVLGSGFRILMCTKVYMDDELTAHENGAIQYDM 383
 Db 334 KRAFDEKRWVSWLPQKRGVAYSIGTKISKYLIMNTNSLRDQIYVHELGHSMHS- 392
 Oy 384 AYAAQPFLLRNGANGFHE---AVGEIMLSAATPKHLKSGILSPDFQEDNETELNFK 439
 Db 393 -----LYSNRQKITYSDYKIFYAIIASIS-----NEVYLNLYYL 425
 Oy 440 -----LKALITVGLTPTTYMLEKRWVVFGEIPKDOMKKMKMKREI 479
 Db 426 LEKYKNDLEKMLTLDKMSIGFAITTRQYITISNFEMWILNLSGAPPTADVYKRYOK 485
 Oy 480 MKREIVG--VVEP-----VPHDEYCDPASLFHVSNDYSFIRYTYRTLYQF 523

Db 486 LELETKPIVEDLNSIYSSTIPLRIPH-----FYGNFY-----YKK 526
 QY 524 QFOALCOAKHHEGFLHKDISNSTEAGOKLFNMLRLG-----KSEPW 566
 Db 527 AVGVAAIISGHR-----VETKVTGAKKQVPELSSGSKDPLDTIKLGVDTLRQAW 580
 QY 567 TLALENVGAKNMNVRPLNVEPLFTWLDK 597
 Db 581 QEALEIVK-----LWIKD 593

RESULT 19

zinc metalloproteinase oligoendopeptidase F U0521 [imported] - Ureaplasma urealyticum
 C:Species: Ureaplasma urealyticum
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C:Accession: D82881
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
 submitted to Genbank, February 2000
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
 A:Reference number: A82870
 A:Accession: D82881
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-611 <GLA>
 A:Cross-references: GB:AE002150; GB:AF222894; NID:96899515; PIDN:AAF30934.1; GSPDB:GN001
 A:Experimental source: serovar 3; biovar 1
 C:Genetics:
 A:Gene: pepF-2; U0521
 A:Genetic code: SGC3

Query Match 3.2%; Score 136; DB 2; Length 611;
 Best Local Similarity 19.5%; Pred. No. 0.2;

Matches 119; Conservative 100; Mismatches 256; Indels 134; Gaps 27;

QY 49 NYNNNTEENYQNNMNGDKSAFLK-EQSTLAQMYPLQELQNTYK----- 94
 Db 69 NYVSNKLTQTNL-IDNEMLAWSQKIEHQHRAKIF--INFENLAIKKDLINSYLNSS 124
 QY 95 -----LQALQONGSSVLSSEKSKRLNTILNTMTI-----YSTG 130
 Db 125 LIKQLEFEELMREKHIHLNQOKVTAISRFSSEGDIFDVLDSMQYQDSINKKQ 184
 QY 131 KVCNPNDQECLELPGINETMANSLDYNERLWAMESKRSVEGQLRPLVEEYVKNEM 190
 Db 185 KVCERN-QTDLVYASKSNDRALKRSAYESHFKAIYDLRNTFSKL--LYEY-VKQNEL 238
 QY 191 ARAHHEDYGDYWGDEYVNGVDYDSRQGLIDVEHTEFEIKRPLEYHLAAVYR-AKL 248
 Db 239 AKLHNFKY-----ISDAFSPKVDKNF-----INHITQTKKRAKG 275
 QY 249 MNAAPSYISPTGCLPAHLIGDMGRFWNTNLSLTVPFGQKFNIDVTAM-VDAQMDAORI 307
 Db 276 INRTKYRT-----LFLKQYQLKRVPEWKNLIIIDKKMFSIESAKNL 320
 QY 308 EKREKFEVSVGLPMTQGFWEKSMIDPQVQKAVCHPTAWDIDGKDFRILMCTKYTMD 367
 Db 321 TLEAALLGSEYINVQAFNEQWISMPNNKISGAVSISNTKGLDKIFILMYDETYN 380
 QY 368 DFLFHHMGHIQDMAAOPFLRLRNGANGFEAVGEISLSAATPKHLKSLGLSPD 427
 Db 381 SLTIVLHGLSHVH--TFFA-----NOSQEVNMEYEFYAEISITNEILMNHHLK-K 431
 QY 428 FOEDNETEINFLKQALITVGLPTTYLLEKRMWVFEKGLPKDQMKKWMEMK-REIVG 486
 Db 432 YENDOLMLRYLIDEMISGFIATTTROAIFSNFEWVA-----NEMINQGEFSMNKIVL 484
 QY 487 VVEPVPHEET-----YCDPASLFHNSNDYSFIRYTRTLRYQRFQ-VALCOAK 534
 Db 485 AYLEINHDYTGKYNKKNISKYDEANALILNIPH---FYTGNYKYYKVIQIC--- 536
 QY 535 HEGPLHKCDI--SNTEAGOKLFNMLRLGK-EPWTLALENVGAKNMNVRPLNVEPLF 592

Db 537 --GLINAIIRFNKNAKEKYCFEKKSGSLSP--LEFINILDIK-INENQWEEVNTIIF 591
 QY 593 -TWLKDQNK 600
 Db 592 NSMIDYIK 600

RESULT 20

carboxypeptidase homolog ypwA - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: D69943

R:Kunst, F.; Ogasawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Be-

C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.;

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Ga-

leher, J.; Harwood, C.R.; Hensaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo,

Koehler, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinc

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Ma-

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowaka, A.; Se-

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpestre, P.; Tognoni, A.; Tosato, V.; Uchiy

T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Dancilin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: D69943

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-501 <KUN>

A:Cross-references: GB:Z99115; GB:AL009126; NID:92634478; PIDN:CAB14125.1; PID:926344

A:Experimental source: strain 168

C:Genetics:

A:Gene: ypwA

C:Superfamily: Thermus aquaticus carboxypeptidase Tag

Query Match 3.1%; Score 135; DB 2; Length 501;
 Best Local Similarity 20.9%; Pred. No. 0.18;

Matches 98; Conservative 65; Mismatches 164; Indels 142; Gaps 26;

QY 180 YEEYVYL-KNEAA-RAHHEDYG-----DYMGQDEYVNGVGYDS 218
 Db 103 YKEVYILCSAETAWBEAKGSDPSLSPYLEQLEFNKRFTTYW-----GYQ-- 150
 QY 219 RGQILDEVEHTEFEIKRPLEYHLAAVYRAKLMAVPSYISPTGCLPAHLIGDMGRFWNTL 278
 Db 151 -----EHPYALDLDFE--PGYTVAYLD-----QLFAEL 177
 QY 279 YSLTVPF-----GQPNIDVTAMVDQAQRIEKAERFVSVGLPMTQGFWEKNS 331
 Db 178 KEALIPLVKQVTAASKNRP-----DTSFTKAPKKEKOKELSLYFQELGYF--- 224
 QY 332 MLDDPGVQKAVQHPPTAWMDGKDFRILMCTKYTMDPLTA-----HHMGHIQYD--MAY 385
 Db 225 --DGGMLDFTV--HPRATTLNKGDAVR--TTRIDKDFRALTGTTHGCHALYEONIDE 278
 QY 386 AAFPFLRNGANGFHEA-----VGE-----IMSLAATPKHLKSLGLSPD- 428
 Db 279 ALSGTNSDASANGIHSSQSLFEENFGRNKKHWTYKRIQASVPQKDISL--DDEV 336
 QY 429 QEDNETEINFLKQAL-LITVGLPTTYLLEKRMWVFEKGLPKDQMKKWMEMKREIVG 486
 Db 337 RAINESKPSPIRVEADELFTPLHIIIRYEIEK---AIFSEVSEDELPSLIMOKKQDYLG 393
 QY 487 VVEPVPHEETYPASLFHNSNDYS-FIRYTRTLRYQRFQ-VALCOAK 534
 Db 394 I---TPQTDAGLIDVHWAGGDFGFPSTALCYMTAAQLKQMLDELEFQALLERGF 450
 QY 540 HKCDISNSTEAGOKLFNMLRLG-KSEPWTLALENVGAKNMNVRPLNVEPLF 587

Fri Mar 14 10:00:50 2003

us-09-978-385-2.rpr

Page 12

Db 451 HPK-----OWLEKVVHIGKRRKKPDIDIKDATG-BELNVRYLIDY 490

Search completed: March 13, 2003, 16:58:59
Job time : 33 secs

Fri Mar 14 10:00:51 2003

us-09-978-385-2.rsp

Page 1

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 16:53:22 ; Search time 15 Seconds
(without alignments)
2225.896 Million cell updates/sec

Title: US-09-978-385-2
Perfect score: 4291
Sequence: 1 MSSSSWLLSLVAVTAOST.....ISKGNPFGONTDDVOTSF 805

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	31.3	732	1 ACET_HUMAN	P12866 homo sapien
2	1337	31.2	1306	1 ACET_HUMAN	P12861 homo sapien
3	1334	31.1	732	1 ACET_MOUSE	P22867 mus musculu
4	1334	31.1	1312	1 ACET_MOUSE	P09470 mus musculu
5	1312	30.6	1193	1 ACET_CHICK	Q10751 gallus gall
6	1310	30.5	1313	1 ACET_RAT	P12822 rattus norv
7	1283.5	28.9	1310	1 ACET_RABIT	P12868 oryctolagus
8	1283	25.4	615	1 ACET_PROMIE	Q10714 drosophila
9	1058	24.7	611	1 ACET_HAIE	Q10715 haematobia
10	1058	24.7	501	1 ACET_HAIE	P50848 bacillus su
11	135	3.1	501	1 ACET_HAIE	P39118 bacillus su
12	135	2.9	627	1 ACET_HAIE	P30195 staphylococ
13	133	2.9	986	1 ACET_HAIE	P30195 staphylococ
14	121	2.8	1034	1 ACET_HAIE	P30195 staphylococ
15	119.5	2.8	1283	1 ACET_HAIE	P30195 staphylococ
16	119.5	2.8	3655	1 ACET_HAIE	P30195 staphylococ
17	119.5	2.8	4540	1 ACET_HAIE	P30195 staphylococ
18	118	2.7	663	1 ACET_HAIE	P30195 staphylococ
19	118	2.7	3911	1 ACET_HAIE	P30195 staphylococ
20	115.5	2.7	3433	1 ACET_HAIE	P30195 staphylococ
21	115	2.7	950	1 ACET_HAIE	P30195 staphylococ
22	114.5	2.7	901	1 ACET_HAIE	P30195 staphylococ
23	114.5	2.7	1225	1 ACET_HAIE	P30195 staphylococ
24	114	2.7	1284	1 ACET_HAIE	P30195 staphylococ
25	113.5	2.6	6669	1 ACET_HAIE	P30195 staphylococ
26	112.5	2.6	773	1 ACET_HAIE	P30195 staphylococ
27	109.5	2.6	1398	1 ACET_HAIE	P30195 staphylococ
28	109	2.5	906	1 ACET_HAIE	P30195 staphylococ
29	108.5	2.5	1084	1 ACET_HAIE	P30195 staphylococ
30	108.5	2.5	2167	1 ACET_HAIE	P30195 staphylococ
31	107.5	2.5	1689	1 ACET_HAIE	P30195 staphylococ
32	107.5	2.5	3685	1 ACET_HAIE	P30195 staphylococ
33	107	2.5	736	1 ACET_HAIE	P30195 staphylococ

RESULT 1	ACET_HUMAN	STANDARD:	PRT:	732 AA.
34	106.5	2.5	1312	1 RASO_YEAST
35	106	2.5	1279	1 BCBH_CHLVI
36	105.5	2.5	645	1 REP_BUCAI
37	105.5	2.5	1002	1 POL_HVI04
38	104.5	2.4	757	1 RTSL_YEAST
39	104	2.4	639	1 GIGB_BACST
40	104	2.4	834	1 CASL_HUMAN
41	103.5	2.4	569	1 CTSP_PLAFA
42	103	2.4	511	1 CTAQ_THBAQ
43	103	2.4	1162	1 BXEN_CLOBU
44	102.5	2.4	660	1 MM02_HUMAN
45	102.5	2.4	704	1 GIGB_YEAST

ALIGNMENTS

ACET_HUMAN
ID ACET_HUMAN STANDARD: PRT: 732 AA.
AC P22966; 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Angiotensin-converting enzyme, testis-specific isoform precursor
DE (EC 3.4.15.1) (ACE-T) (dipeptidyl carboxypeptidase I) (kininase II).
GN DCP1 OR DCP OR ACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=9004671; PubMed=2554286;
RX Ehlers M.W., Fox E.A., Strydom D.J., Riordan J.F.;
RA "Molecular cloning of human testicular angiotensin-converting enzyme;
RT the testis isozyme is identical to the C-terminal half of endothelial
RT angiotensin-converting enzyme."
RT Proc. Natl. Acad. Sci. U.S.A. 86:7741-7745(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89338720; PubMed=2547653;
RA Lattion A.L., Soubrier F., Allegrini J., Hubert C., Corvol P.;
RA "Alteme-gelase F":
RT "The testicular transcript of the angiotensin I-converting enzyme
RT encodes for the ancestral, non-duplicated form of the enzyme."
RN FEBS Lett. 252:99-104(1989).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANTS P-32; G-49 AND S-712.
RX MEDLINE=9251580; PubMed=10319862;
RA Rieder M.J., Taylor S.L., Clark A.G., Nickerson D.A.;
RA "Sequence variation in the human angiotensin converting enzyme."
RN Nat. genet. 22:59-62(1999).
RN [4]
RP ZINC-BINDING;
RX MEDLINE=91308093; PubMed=1649623;
RA Ehlers M.R., Riordan J.F.;
RA "Angiotensin-converting enzyme: zinc- and inhibitor-binding
RT stoichiometries of the somatic and testis isozymes."
RN Biochemistry 30:7118-7126(1991).
RN [5]
RP BIOCHEMISTRY 30:7118-7126(1991).
RT "FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
RT THE C-TERMINAL HIS-LEU. THIS RESULTS IN AN INCREASE OF THE
RT VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN. ALSO ABLE TO INACTIVATE
RT BRADYKININ, A POTENT VASODILATOR."
RN [6]
RP CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
RN oligopeptide-l-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
RN Asp nor Glu. Converts angiotensin I to angiotensin II.
RN [7]
RP CORRELATION: BINDS 1 ZINC ION.
RN [8]
RP SUBCELLULAR LOCATION: Type I membrane protein.
RN [9]
RP ALTERNATIVE PRODUCTS: 2 isoforms, testis-specific (shown here) and
RN alternative (AC P12821), are produced by alternative splicing.
RN [10]
RP TISSUE SPECIFICITY: SPERMATOCYTES, ADULT TESTIS.
RN [11]
RP INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL

FT	VARIANT	361	361		/ftid=VAR_011708.
FT	VARIANT	1286	1286	R -> S.	/ftid=VAR_011709.
FT	CONFLICT	35	35	O -> E (IN REF. 2).	D -> R (IN REF. 2).
FT	CONFLICT	42	42		
SO	SOURCE	1306 AA;	149714 MM;	1B33BCA7301A26AA	CRC64;
	Query Match		31.2%;	Score 1337;	DB 1;
	Best local similarity	41.7%;	Pred. No. 9.4e-86;		
	Matches	255;	Conservative 118;	Mismatches 204;	Indels 34;
				Gaps	
QY	20	TIEGQAKTEFLDKRHNHAEDELFOSSLASNVNTNTEE-----VWQMNNAGDWKSA	71		
Db	644	TDEEAKFEVEEDRRSQVVMNEYAFANNNTYNITTTTSKLILKKMMQLANHF-----	697		
QY	72	FLKEOSTLAQMYPLBOELONLTVKLQALQOOGSSVLSEDSKSRLNTLLTMSIYSTGK	131		
Db	698	--LKYGQAARFVNLOLTMTIKRIKKQODLERALBAQOELEENKILLDMETYSVAT	755		
QY	132	VGNNDNGOELLLEPGNEIMANSLDNYRLAMESMWSREVGKOLRPLEYEVYLKNEMA	191		
Db	756	VCHNG--SCQLPEPDLTNMAISRKYEIDLLMANGMDKKGRALLQFPYYVELINOAA	813		
QY	192	RANHVEDGYDRWDGEVNGVDGDYDSRGQLIEDVBHTPEEKLPYLEHLAYVAKLMA	251		
Db	814	RLNGYDAGOSWRSMYETPSLE-----QDLERLFEOLEDPLYLNLHAHYRRALHRH	863		
QY	252	Y-PETISPICLRPHLLIGDMMGREFMTNLSTVPFGOKRNIDVTPAWDQAMDQRIFE	310		
Db	864	YGAOHINLEGIPAHLLIGNMAQTWSNIYDLVPPPSAPSDYTTPAMLKGCTPRRMFE	923		
QY	311	AEEFVSGLPNMTIOGFEMSNMLTDGNAVOKACVCEPTAMDLGK--DFRILMCTVYMDF	369		
Db	924	ADDEFSLTGILRPVRPERWNKSMLEKFPDGREVVCHASANDFYNGKDFXICTTYVNEDL	983		
QY	370	LTAHHMEGHIDYMAIYAOPPLRLNANGANGFNHANGEIMSIISAAPPKLSIGLSPDQ	429		
Db	984	VVAHEMKHIQFYAQYKDLPVALREGANPGEFHAIIDVALSVTPKHLSINLLSSGG	1043		
QY	430	EENEIEIFELKOALTIVGTLPETYMLEKRWAVFEGEILPKDOMKKWMENKRELIVYE	489		
Db	1044	SD-EHDINFELKMALDKIAFLPEYLIYLDQRMHVDPGSTITKENYNOEWMSLTKYQGLCP	1102		
QY	490	PVPHDETYPASLEFHSVNSDISFRITYTRTLPLYFOQEALCOAKHEGPLHCIDSNSTE	549		
Db	1103	PVPRTIOGEDFGAKFHLPSSVPIRYRVSFIIQOFHEALCOAAGHTGPLHKCDIYOSKE	1162		
QY	550	AOCKLEFMNLRLKGSPEPTALEENVAGKANNNVPLNTPYEPELTWLKDNK--NSFWGM	606		
Db	1163	AOORLATPAKRGFSWPMEAMOLITGOPNMNASAMLSYFRPLDMLRTNELHGKCLMP	1222		
QY	607	SDWSPSYADS 617			
Db	1223	QYNWTPNSARS 1233			
RESULT 3					
ACET_MOUSE					
ID ACET_MOUSE	STANDARD;	PRT;	732 AA.		
AC P22967:					
DT 01-AUG-1991 (Rel. 19,	Created)				
DT 01-AUG-1991 (rel. 19,	Last sequence update)				
DT 15-JUN-2002 (Rel. 41,	Last annotation update)				
DE Angiotensin-converting enzyme, testis-specific isoform precursor					
(EC 3.4.15.1) (ACE-r) (Dipeptidyl carboxypeptidase I) (Kininase II).					
DE DCPI OR ACE.					
OS Mus musculus (Mouse).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OX NCBI_TaxId=10090;					
RN [1]					
RP SEQUENCE FROM N.A.					

```

RX MEDLINE-90318396; PubMed-2164636;
RT Howard T.E., Shal S.-Y., Langford K.G., Martin B.M., Bernstein K.E.;
RT "Transcription of testicular angiotensin-converting enzyme (ACE) is
RT initiated within the 12th intron of the somatic ACE gene."
RL Mol. Cell. Biol. 10:4294-4302(1990).
CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide-1-xaa-xbb, when xaa is not pro, and xbb is neither
CC asp nor glu. Converts angiotensin I to angiotensin II.
CC -1- COFACTOR: BINDS 1 ZINC ION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; testis-specific (shown here) and
CC somatic (AC P09470); are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: SPERMATOCTYES, ADULT TESTIS.
CC -1- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL
CC REGULATION BY ANDROGENS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL; M55333; AAA37149.1; -
DR EMBL; M61094; AAA37150.1; -
DR PIR; A35655; A35655.
DR MEROPS; M02.004; -.
DR MGD; MGI:87874; Ace.
DR InterPro; IPR001348; Peptidase_M2.
DR InterPro; IPR000130; Zn_MTpeptidse.
DR Pfam; PF01401; Peptidase_M2; 1.
DR PRINTS; PR00791; PEPTIDASEA.
DR PRODOM; PD004184; Peptidase_M2; 1.
DR PROSITE; PS00142; ZINC_PROTASE; 1.
KW Hydrolyase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
KW Glycoprotein; Transmembrane; Testis; Signal; Alternative splicing.
FT CHAIN 1 31
FT SIGNAL 1 31
FT POTENTIAL.
FT ANGIOTENSIN-CONVERTING ENZYME,
FT TESTIS-SPECIFIC ISOPFORM.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT ZINC (CATALYTIC) (BY SIMILARITY).
FT BY SIMILARITY.
FT ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 413 413
FT ACT_SITE 414 414
FT METAL 417 417
FT METAL 441 441
FT METAL 441 441
FT CARBOHYD 102 102
FT CARBOHYD 120 120
FT CARBOHYD 139 139
FT CARBOHYD 139 139
FT CARBOHYD 185 185
FT CARBOHYD 367 367
FT CARBOHYD 616 616
FT SEQUENCE 732 AA; 84047 MW; 16C81E7FBD09BD9 CRC64;
Query Match 31.1%; Score 1334; DB 1; Length 732;
Best Local Similarity 42.6%; Pred. No. 6,7e-86;
Matches 253; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

```

```

Db 187 TCMPEPLTLNMMATSRKYEELLMWAKSMWRDVGRAILPEPPKVERSNKIAKNGYTD 246
Qy 200 GDWVGDEYVNGVDYDSRGQLLEDVHFEELKPIYEHHAIVAKLNATPS-YISP 258
Db 247 GDSWMSLYESDLE-----QDLKTYOELDPLLYLHLHYVRSLSLRHNGSEYIML 296
Qy 259 IGLCPAHLGLDMGWFNTLYSLVFPFGOKPNIDVTAMVDQMDQRIKREKREFFVY 318
Db 297 DGPILPAHLGLMMQWTSNIYDLVAPPSPAPNIDATAMKQWTRRIKFEKDNFTSL 356
Qy 319 GLPNNTOGFEMNSMLTDPGNVQKAVCPRTAMDJGK-DPFIIMCTVYTMDFLTAHENG 377
Db 357 GLPVPPEFWKMSLEKPTGREGVCHPSAMDYNGDFRIKQCTSVNMDLVIAHENG 416
Qy 378 HIOYDMAAOPFLRLNAGCEFEHAEVGEIMSLSAATPKRLKSTGLSPDFEDNETFN 437
Db 417 HIOYFMQKDLVPYFREGANPGEHAGDIMALSVTPKHLVSLNLSTE-GSGYEYDIN 475
Qy 438 FLKQALTYGTLPFTYMLERKMWVFKGEIPDKMMKRWEMKREIVGVPEVPHDETY 497
Db 476 FLKMAALDKIAFIPFSYLIDQWWRVFDGSIKENYNQEWMSLRKYGQICPPVPRSGD 535
Qy 498 CDPASLPHVNDYSFIRYTRTLQFOFQALQAKHSGPLHKCDINSTAGCKLFNM 557
Db 536 FDPGSKRHVAVNPYVYFVSFIIOFHQALCRAAGHGPLHKCDIYQSKRAGKLADA 595
Qy 558 LRIGKSEPTWLTALENVVAGKANMVRPLNYFEPFLTKDKQK--NSFVGM-STDMSP 612
Db 596 MKLGYSKPWEAMKLTIGOPNMSASAMNMFKLTPEMLTEBRRRGETLGMWEYIMAP 653

RESULT 4
ACE_MOUSE
ID ACE_MOUSE STANDARD: PRT; 1312 AA.
AC P09470;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Angiotensin-converting enzyme somatic isoform precursor (EC 3.4.15.1)
DE (ACE) (Dipeptidyl carboxypeptidase I) (Mininase II).
GN DCP1 OR ACE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89308599; PubMed-2545691;
RX Bernstein K.E., Martin B.M., Edwards A.S., Bernstein E.A.;
RX "Mouse angiotensin-converting enzyme is a protein composed of two
RX homologous domains."
RL J. Biol. Chem. 264:11945-11951(1989).
RL [2]
RP SEQUENCE OF 1-332 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-88298730; PubMed-2841312;
RX Bernstein K., Martin B.M., Bernstein E.A., Linton J., Striker L.,
RX Striker G.;
RX "The isolation of angiotensin-converting enzyme cDNA."
RL J. Biol. Chem. 263:11021-11024(1988).
CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide-1-xaa-xbb, when xaa is not pro, and xbb is neither
CC asp nor glu. Converts angiotensin I to angiotensin II.
CC -1- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; somatic (shown here) and testis-
CC specific (AC P22967); are produced by alternative splicing.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

[illegible]

```

RESULT 5
AC CE CHICK
ID ACE_CHICK STANDARD; PRT; 1193 AA.
AC 010751;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Angiotensin-converting enzyme (EC 3.4.15.1) (Dipeptidyl
DE carboxypeptidase I) (kininase II) (Fragment).
GN DCP1 OR ACE.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
CX NCBI_TaxID=9031;
XX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Lung;
RC MEDLINE=95110342; Pubmed=7811282;
RX Escher C.R., Thomas K.E., Bernstein K.E.;
RA "Chicken lacks the testis specific isozyme of angiotensin converting
RT enzyme found in mammals.";
RL Biochem. Biophys. Res. Commun. 205:1916-1921(1994).
CC -I FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
CC -I CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide-I-xaa-xbb, when xaa is not Pro, and xbb is neither
CC Asp nor Glu. Converts angiotensin I to angiotensin II.
CC -I COFACTOR: BINDS 2 ZINC IONS.
CC -I SUBCELLULAR LOCATION: Type I membrane protein.
CC -I SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collabor
CC between the Swiss Institute of Bioinformatics and the EMBL outstat
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no
CC modified and this statement is not removed. Usage by and for comme
CC entities requires a license agreement (See http://www.isb-sib.ch/anno
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, LA0175; AAA75554.1; -.
DR MEROPS; M02.001; -.
DR InterPro; IPR001548; Peptidase_M2.
DR InterPro; IPR00130; Zn_MPeptidase.
DR Pfam; PF01401; Peptidase_M2; 2.
DR Prodom; PD004184; Peptidase_M2; 2.
DR PROSITE; PS001442; ZINC_PROTEASE; 2.

```

KW Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 KW Glycoprotein; Transmembrane; Repeat.
 FT NON TER 1 1156 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1157 1173 POTENTIAL.
 FT TRANSMEM 1174 1193 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 125 481
 FT REPEAT 723 1079
 FT METAL 288 288
 FT ACT SITE 289 289
 FT METAL 292 292
 FT METAL 886 886
 FT ACT SITE 887 887
 FT METAL 890 890
 FT CARBOHYD 11 11
 FT CARBOHYD 60 60
 FT CARBOHYD 216 216
 FT CARBOHYD 407 407
 FT CARBOHYD 447 447
 FT CARBOHYD 485 485
 FT CARBOHYD 513 513
 FT CARBOHYD 555 555
 FT CARBOHYD 575 575
 FT CARBOHYD 658 658
 FT CARBOHYD 1089 1089
 SQ SEQUENCE 1193 AA; 137820 MW; 954472A18EA471C7 CRC64;

Query Match 30.6%; Score 1312; DB 1; Length 1193;
 Best Local Similarity 40.4%; Pred. No. 4.7e-84;
 Matches 264; Conservative 110; Mismatches 243; Indels 36; Gaps 10;

QY 22 EDQAKFLDKFNHEDLFYQSSLASWNTNTEENVQNMNAGDKSAFLKEOSTLAQ 81
 DB 544 EQAAKEFLSEYNSTAEVNNATTEASWEYNTNITDHNEKVMLEKMLAMSKHTEYGMAR 603
 QY 82 MYPLGEIONLYKLOLQALQNGSSVLSSEKSKRNTLITNTSTYSGKVCNPN--P 138
 DB 604 QPDBDFODEYTRILINKLSYERALPEDELKEYNTLSDMETTYSAKVCARENTEHP 663
 QY 139 QECLELPEGLNIMNSLDYNERLWMSWSEVQKRLPYEEYVLKNEMARANYED 198
 DB 664 ----LDPLDILATSRDYNELLPAMKQWMDASGAKIKDKKRYVELSNKAAYNGTGD 718
 QY 199 YGDYRGDYEVNGVGYDYSKQOLIEDVHPEETKPYEHLHAYVRAKLMNAY-PSYIS 257
 DB 719 NAYVRSLEYETPTFE-----EDLERLYLQLOPLYLMDHAYVRALNKKYGAHIS 768
 QY 258 PIGCLPAHLGDMGREFWNTLSLTPFGQKPNIDYDAMVDQMDAQRIKEAKPPVVS 317
 DB 769 LKGPILPAHLGMMMAQSNIFDLMPDPDAKYDAPPAKMQOQWTPKPMFEESDREFTS 828
 QY 318 VGLPMGTQGEWNSMLTDPGNVAKAVCHPTAMD-L-GKGFRLILCTKYTMDFLTAHHEM 376
 DB 829 LGLIPMPQEFMDKSWIEKPADGREVYCHASAMDFNKRDFIKICTYVNMMDLITVHEM 888
 QY 377 GHIOYDAMAYAOPLRLRGANGEGHEANGELMSIATPKLKSIGLISPDQEDNETEI 436
 DB 889 GHVOFLQYMDQPLSFRGANGGEHALGDYALSVSPKHLISINLLD-QVTEHESDI 947
 QY 437 NEFLKQALTIYGLTPTMYLKKRMWYKGEIPKDDMKWKKWEMKRELVGVGVPEVDET 496
 DB 948 NYLMSIALDKIAFLPFGILMDQMKKVPDGIKIKEDYNOQMMNRLYYQIGCLPVPVSED 1007
 QY 497 YCDPASLHVSNDSYFLIRYTRILYOFQFQALCOAAKHESPLHKDISNSTENGOKLFFN 556
 DB 1008 DEDGAKFHIANPYIRIEYFVIOFQFQALCKAAGHGPHLTGCIISKEGSKGLGD 1067
 QY 557 MLRIKSPWPLALENVGAKNNVRLPLNYFEPLFWL--KQONKSFYGM-STDMSPY 613
 DB 1068 AMKLGFSKPEWPAOLLITGQNMSEAELMSYFELMTLVKKNTEENGEVGLWPEYSWTPY 1127
 QY 614 ADOSIKVRIISALG-----DKAYEMNDNEWYLFRRSSVAAVAMQYELKVK 659

DB 1128 ATEFHAAITADFLGMSVGTQKATAGAW-----VLALALVFLTITSIFLGK 1175
 RESULT 6
 ACE_RAT
 ID ACE_RAT STANDARD: PRT: 1313 AA.
 AC P47820.
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)
 DE (ACE) (Dipeptidyl carboxypeptidase I) (kininase II).
 GN DCP1 OR ACE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_Taxid=10116;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=94121658; PubMed=8292044;
 RA Koike G., Krieger J.E., Jacob H.J., Mukoyama M., Pratt R.E.,
 RA Dzeu V.J.;
 RT "Angiotensin converting enzyme and genetic hypertension: cloning of
 RT rat cDNAs and characterization of the enzyme";
 RL Biochem. Biophys. Res. Commun. 198;380-386(1994).
 RN {2}
 RP SEQUENCE FROM N.A.
 RC STRAIN=LBN/N; TISSUE=Lung;
 RA Jafarian-Tehrani M., Listwak S., Barrientos R.M., Michaud A.,
 RA Corvol P., Sternberg E.M.;
 RT "Characterization of a missense mutation in the angiotensin
 RT I-converting enzyme cDNA in exudative inflammation resistant F344/N
 RT rats";
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU. THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-I-xaa-xbb, when xaa is not pro, and xbb is neither
 CC asp nor glu. Converts angiotensin I to angiotensin II.
 CC -1- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: somatic (shown here) and testis-
 CC specific; are produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC -----
 CC EMBL: U03708; AAA82110.1; -
 CC EMBL: U03734; AAA82111.1; -
 CC EMBL: AF201332; AAG35597.1; -
 CC MEROPS: M02.001; -
 CC DR MEROPS: M02.004; -
 CC DR InterPro: IPR001548; Peptidase_M2.
 CC DR InterPro: IPR000130; zn_Mrpeptidase.
 CC DR Pfam: PF01401; Peptidase_M2; 2.
 CC DR PRINTS: PR00791; PEPDIPYASER.
 CC DR ProDom: PD004184; Peptidase_M2; 2.
 CC DR PROSITE: PS00142; ZINC_PROTEASE; 2.
 CC DR Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 CC KW Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.
 CC FT SIGNAL 36 1313
 CC CHAIN 36 1313
 CC FT
 CC FT DOMAIN 36 1265
 CC FT TRANSMEM 1266 1282
 CC POTENTIAL.
 CC EXTRACELLULAR (POTENTIAL).
 CC ISOFORM.

```

FT DOMAIN 1283 1313 CYTOPLASMIC (POTENTIAL).
FT REPEAT 233 589
FT REPEAT 831 1187 ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT METAL 396 396 ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT ACT SITE 397 397 ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT METAL 400 400 ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT METAL 994 994 ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT ACT SITE 995 995 ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT METAL 998 998 ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 683 683 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 701 701 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 766 766 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 948 948 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1197 1197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 207 207 R -> K.
SQ SEQUENCE 1313 AA; 150907 MW; 8CB5D0015F129591 CRC64;

Query Match 30.5%; Score 1310; DB 1; Length 1313;
Best Local Similarity 42.0%; Pred. No. 7, 4e-84;
Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps 7;

QY 20 TIEQAKTEFLDKNEHAEDELTYOSSLASWNTNTINEVONNNANNAKWSAFLEKQSTL 79
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 650 TDEAKANRFEVEYDRAKAVLMNEYAEAMNHYNTNITIESSKILLQNKKEYSNTTLKYGW 709
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 80 AQMYPLQEIOLNLFVKLQALQNGSSVYSEDKSKRLNTLTMTSTYTGKVCNDNQ 139
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 710 AKTEFVSNQDNSTIKRIKKVQNVDAVLPNPLEEYNOILLMETTYAVANCYNG-- 767
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 140 ECLLEPGLEINANSIDYNERLWAMESRSEYKOLRPLYEYVYLKMKMARANYEY 199
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 768 TCSLEPDLNIMATSKREBELLMWKSMDKAGRLFFPKYVDFSKTALNLGSDA 827
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 200 GDYWRDYEYNGVDYDYSRGQLEDEYHEETFEIKPLYEHLNAYRAKLNNAPPS-YISP 258
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 828 GDSMRSSYESDLE-----QOLEKLYQELQPLVLYNLHAYVRSLSHRHGSXYNL 877
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 259 ICGLPRLHLLGDMGRFWTNLSTLTFPGQKPNIDVDAMVDAQMDAQRIFKEAEKFFSV 318
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 878 DGPRLPHLLGNMMAQWTSNITYDLVAFPPSPASIDATEAMIKQGWTPRRIFKEADNFTSL 937
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 319 GLPNMTQGEFENSMFLDPGNVOKAVCHPTAMDLGK-DPRILMCTKVYMDDFLAHENG 377
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 938 GLLPYRPEFENKSMLEKPTDGRVYVCHASMDPYNKDRIRKOCISVNMEEVIAHHENG 997
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 378 HIQYDMAVAQAQPLRLNGANEGFHEAVGEIMLSAATPKHLISIGLSDFQEDNTEIN 437
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 998 HIQYFQKXLPYTFREGANPGFHEAIGVLAIVSTPKHLISLISE-SGSGYHDIN 1056
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 438 FLKQALTIQGLPETYMLEKRWMEKGEIPKDOMKMKWMEKRAIVGVVEYVDEY 497
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 1057 FLKMKMLDKIAIPFSITLIDQWRVFDSTIKENINQEWMSRLKLYQGLCPVPRSSQD 1116
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 498 CDPAFLHVSNDYSFTRYTYRTLYQFOEALCOAKHKGPLKDCISNSTEAGOKLFW 557
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 1117 FDEGSKFHPANVPYIRYFISFLIOGFHEALCRAAGHGPLKDCIYOSKEGKLADA 1176
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 558 LRIQKSEPTWLALENVAGAKNNANVRLNVEFLFTWLKDQK--NSFGW-STDWP 612
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 1177 MKLGSKQWPEAMKITGQPNMSASAIMNYFKPLTEMLVTENRRHGETLIGMEYTWTP 1234
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

ID ACE_RABBIT STANDARD; PRT; 1310 AA.
AC P12822; C02852;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Angiotensin-converting enzyme, somatic isoform precursor (BC 3.4.15.1)
DE (ACE) (Dipeptidyl carboxypeptidase I) (Kinase II).
GN DCPI OR ACE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA MEDLINE=92178960; PubMed=1311831;
RT "Thekkumkara T.J., Livingston W. III, Kumar R.S., Sen G.C.;
RT "Use of alternative polyadenylation sites for tissue-specific
RT transcription of two angiotensin-converting enzyme mRNAs.";
RL Nucleic Acids Res. 20:683-687(1992).
RN [2]
RP REVISIONS.
RA Sen G.C.;
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-88 FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=91139683; PubMed=1847388;
RT Kumar R.S., Thekkumkara T.J., Sen G.C.;
RT "The mRNAs encoding the two angiotensin-converting isozymes are
RT transcribed from the same gene by a tissue-specific choice of
RT alternative transcription initiation sites.";
RL J. Biol. Chem. 266:3854-3862(1991).
RN [4]
RP SEQUENCE OF 34-55.
RC TISSUE=Lung;
RA MEDLINE=84051289; PubMed=6314908;
RT Iwata K., Blacher R., Soffer R.L., Lai C.Y.;
RT "Rabbit pulmonary angiotensin-converting enzyme: the NH2-terminal
RT fragment with enzymatic activity and its formation from the native
RT enzyme by NH4OH treatment.";
RL Arch. Biochem. Biophys. 227:188-201(1983).
RN [5]
RP THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
RP VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
RN [6]
RP -I- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
RN oligopeptide--I-xaa-xbb, when xaa is not pro, and xbb is neither
RN asp nor glu. Converts angiotensin I to angiotensin II.
RN -I- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
RN -I- SUBCELLULAR LOCATION: Type I membrane protein.
RN -I- ALTERNATIVE PRODUCTS: 2 isoforms; somatic (shown here) and testis-
RN specific (AC P22968), are produced by alternative splicing.
RN -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
RN -----
RN CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.isb-sib.ch/announce/
RN or send an email to license@isb-sib.ch).
RN -----
RN CC EMBL: X62551; CAA44428.1; -.
RN EMBL: M58579; AAA31151.1; ALT_SEQ.
RN PIR: A23455; A23455.
RN PIR: S35484; S35484.
RN MEROPS: M02.001; -.
RN InterPro: IPR001548; Peptidase_M2.
RN InterPro: IPR000130; Zn_MTpeptidse.
RN Pfam: PF01401; Peptidase_M2; 2.
RN PRINTS: PR00791; PEPTIDTASEA.
RN Prodom: PD004184; Peptidase_M2; 2.

```

RESULT 7
ACE_RABBIT


```

FT CARBOHYD 311 311 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 48 51 WAGV -> GPMR (IN REF. 3).
FT CONFLICT 141 141 C -> S (IN REF. 3).
FT CONFLICT 293 293 A -> G (IN REF. 3).
FT CONFLICT 486 486 S -> T (IN REF. 3).
FT CONFLICT 533 533 V -> M (IN REF. 3).
FT CONFLICT 547 547 R -> A (IN REF. 3).
SQ SEQUENCE 615 AA; 71025 MW; 3583DD5EA/F33CFB CMC64;

Query Match 25.4%; Score 1090; DB 1; Length 615;
Best Local Similarity 36.6%; Pred. No. 6.9e-69;
Matches 224; Conservative 122; Mismatches 244; Indels 22; Gaps 9;

OY 8 ILSTAVATAASTIEQAKTFLDKFNHEADLFYQSSIASNNYNTNTEENVQNMNAGD 67
DB 8 ILATLAVTALVKKEIQAKYELENKRLARTVETAAVAAGSNTIDENKKEKKEISA 67
OY 68 KWSAFLKQSTLAQMYPLQEIQNTLVKQLQALQONSSVLSSEKSKRLNTLNTMTIY 127
DB 68 ELAKFMKAVASDPTKFKFORSYQSEDLKROKALTKRLGYAALPEDDYALDLTSLAMSNE 127
OY 128 STGKVCNPDNDPOEC-LILEPGLNFIANSUDYNERLWAMESRSEVGKQLRPLYEYVL 186
DB 128 AKVAKCDKDKSTCKDLADPEIEEYISKRDEHETAYWRREYDKAGTAVRSOPEYVEL 187
OY 187 KENARAHNYEDYGDYWRGDEYVGVGYDYSRQQLIEDVHFEETIKPLYEHLHAYRA 246
DB 188 NTKAKKLNFDTSGAEAWIDEYE-----DDTEQQLEDI---PADRLPQLQCHGVYRF 237
OY 247 KLIMNAV-PSYISPIGCLPAHLIGDMGRFWNTLSLTPVGQKPNIDVTDAWDQAWDAQ 305
DB 238 RLKRLHGDVAVSETGPIPHMLLGMMMAQOWSEIADIVSPPEKPLVDVSAEMEKOAYTPL 297
OY 306 RIFKEAEKFEVSVGLPNNQTOGFWENSMILTPGYNQKAVCHPTAADLQK-GDFRLMTQTV 364
DB 298 KMFQMGDFLTSMLTKLPQDFMOKSLTEKPTQDRDLVCHASANDFLIDVTRKQCTRV 357
OY 365 TMDPFLAHNHMGHIQYDAAVAPPLLRNANGEGFEAVGELMSAATPKHLKSLGL 424
DB 358 TQDQFLVHHEGLHIQYFLQYOHQPFYVRGANGFHEAVGDLVLSVSPKHLKIGLL 417
OY 425 SPDEQDNTEINFLKQALITVGLPTFTYMLKRMWRNKGKGLPDKQMKKWKMKREL 484
DB 418 K-DYVRDDEARINQLEFLTALDKVLEFPATFMKRYMSLRGEVDANMCAWCAKLADEX 476
OY 485 VGVEVPHDEYCDPASLFEHVSNDYSFIYRTTYLQFOFQALC-QAKH-----EGP 538
DB 477 SGLEPVRVRESEKEDAPAKYHISADVEYLRVLSFIQFQFYKACIKAGQYDPDVELP 536
OY 539 LHKCDISNSTEAGOKLFNMLRLGKSEPTLALNNVGAKMNRPLNLFPELTLKQD 598
DB 537 LDMCDIYGSARAGAAFRHMLSMGASKRMPDLEAFNGERLMSGKALAEYFELRVLMLEAE 596
OY 599 N-KNSFVGWST 608
DB 597 NIKNNVHIGWT 608

RESULT 10
ACE_HAIE
ID ACE_HAIE STANDARD; PRT; 611 AA.
AC Q10715;
DT 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 30-MAY-2000 (rel. 39, Last annotation update)
DE Angiotensin-converting enzyme precursor (EC 3.4.15.1) (Dipeptidyl
carboxypeptidase I) (Kininase II).
GN ACE.
OS Haematobia irritans exigua (Buffalo fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Muscoidea; Muscidae; Haematobia.
NCBI_TaxID=34678;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96215437; Pubmed=8647080;
RA Wilfjels G.L., Fitzgerald C., Gough J., Riding G.A., Elvin C.,
RA Kemp D.J., Willadsen P.;
RT "Cloning and characterisation of angiotensin-converting enzyme from
RT the dipteran species, Haematobia irritans exigua, and its expression
RT in the maturing male reproductive system.";
CC Eur. J. Biochem. 237:414-423(1996).
CC - FUNCTION: INVOLVED IN THE SPECIFIC MATURATION OR DEGRADATION OF A
CC NUMBER OF BIOACTIVE PEPTIDES.
CC - CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
CC Asp nor Glu. Converts angiotensin I to angiotensin II.
CC - COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC - TISSUE SPECIFICITY: EXPRESSED IN THE COMPOUND GANGLION AND IN THE
CC POSTERIOR REGION OF THE MIDGUT.
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: I43965; AAA70427.1; -.
CC DR MEROPS: M02.003; -.
CC DR InterPro: IPR001548; Peptidase_M2.
CC DR InterPro: IPR000130; zn_MTPeptidse.
CC DR Pfam: PF01401; Peptidase_M2; 1.
CC DR PRINTS: PR00791; PEPDIPYASEA.
CC DR ProDom: PD004184; Peptidase_M2; 1.
CC DR PROSITE: PS00142; ZINC_PROTEASE; 1.
CC KW Hydrolase; Metalloprotease; Carboxypeptidase; zinc; Dipeptidase;
CC KW Glycoprotein; Signal.
CC FT SIGNAL 1 17
CC FT CHAIN 1 611
CC FT METAL 367 611 POTENTIAL.
CC FT METAL 367 611 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT ACT_SITE 368 368 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT METAL 371 371 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT CARBOHYD 53 53 N-LINKED (GLCNAC... ) (POTENTIAL).
CC FT CARBOHYD 196 196 N-LINKED (GLCNAC... ) (POTENTIAL).
CC FT CARBOHYD 531 531 N-LINKED (GLCNAC... ) (POTENTIAL).
CC SQ SEQUENCE 611 AA; 70505 MW; A43D6F5A83CB53 CRC64;

Query Match 24.7%; Score 1058; DB 1; Length 611;
Best Local Similarity 37.2%; Pred. No. 1.2e-66;
Matches 223; Conservative 108; Mismatches 246; Indels 22; Gaps 7;

OY 8 ILSTAVATAASTIEQAKTFLDKFNHEADLFYQSSIASNNYNTNTEENVQNMNAGD 67
DB 8 ILATLAVTALVKKEIQAKYELENKRLARTVETAAVAAGSNTIDENKKEKKEISA 67
OY 68 KWSAFLKQSTLAQMYPLQEIQNTLVKQLQALQONSSVLSSEKSKRLNTLNTMTIY 127
DB 68 ENAKFLKEVADIDQKFNMRITGSAVDYRQFSLKSTGYSALEADYELLLEVLSAMSNEF 127
OY 128 STGKVCNPDNDPOEC-LILEPGLNFIANSUDYNERLWAMESRSEVGKQLRPLYEYVL 186
DB 128 AKVAKCDKDKSADKDLSDPEIEEYISKRDEHETAYWRREYDKAGTAVRSOPEYVEL 187
OY 187 KENARAHNYEDYGDYWRGDEYVGVGYDYSRQQLIEDVHFEETIKPLYEHLHAYRA 245
DB 188 NTKAKKLNFDTSGAEAWIDEYE-----DAFEDQLEAFEDIKLYQGVGYR 236
OY 246 KLIMNAV-PSYISPIGCLPAHLIGDMGRFWNTLSLTPVGQKPNIDVTDAWDQAWDA 304
DB 237 YRLKFKYGDVSVKTPGLPHLLGMMMAQOWSEIADIVSPPEKPLVDVSDENVAAGYTP 296
OY 305 QRIKFEAEKFEVSVGLPNNQTOGFWENSMILTPGYNQKAVCHPTAADLQK-KGDFRIIMCTK 363

```

```

Db 297 LKMFQWMDDFEQSGMLKLEQEFWDKSLLEKPPDGRDLVCHASAMDFYLLDDVRIKQCTR 356
Oy 364 VTMDFTLHHEMGHIQYDAMAYAAQPLLRNGANEHGVGIMSLSATPRHHTKISGL 423
Db 357 VTQOQEFVHEHMGHIQYFLOYQHPVYRTPKNGPCHGVGLSLSTFPHLEBVL 416
Oy 424 LSPDFQENEFNEFLKQALTYGLTPFTYMLEKRWVFKGEIPRDMKKWEMKRE 483
Db 417 LK-NVSDNEARINQLETLADKIVFLPFAFTMDKRYMALFROADKSEKNCAPFWLREB 475
Oy 484 IGVVEPVPHDETCDPASPSEFHVSNDSYFIRYTRTLQFOFQALCOAA-----KHBC 537
Db 476 YGIEPPVYRTEKDFDAPAKYHVSADVEYLRIVSFILOFQYKXACITGEVVPNQTEY 535
Oy 538 PLHKCDISNSTEAGOKLFNMLRLKSEFWTTLLENVYAKMMNVRPLNTEPEPLTWL 596
Db 536 PLNDNDIYGSKEAGKLFENMISLGSKFPDPALEAFNGERIMTKALAEFEPLRWLVE 594

RESULT 11
YFWA_BACSU STANDARD: PRT: 501 AA.
AC P50848;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical metalloprotease yPWA (Ec 3.4.24.-).
GN YFWA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fadre C., Ferrari E., Galleron N.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi A.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kunano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Nodack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portolick S., Prescott A.M.,
RA Priescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadele Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror F., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Tosiato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,
RA Viari A., Wambuit R., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RA "The complete genome sequence of the Gram-positive bacterium Bacillus
RT

```

```

RT subtilis";
RL Nature 390:249-256(1997)
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M32.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC
CC EMBL: L47838; AAB38482.1; -
CC EMBL: L77246; AAA96610.1; -
CC EMBL: Z99115; CAB1425.1; -
CC MEROPS: M32.0PW.
CC Subtilisin; B611458; yPWA.
CC Interpro: IPR001333; Peptidase_M32.
CC Interpro: IPR00130; Zn_MTPeptidse.
CC Pfam: PF02074; Peptidase_M32; 1.
CC PRINTS: PR00998; CRBOXYPTASE.
CC PROSITE: PS00142; ZINC_PROTEASE; FALSE_NEG.
CC Hypothetical protein; Hydrolase; Metalloprotease; Zinc;
CC Complete proteome.
CC METAL 265
CC ACT_SITE 266
CC METAL 269
CC SEQUENCE 501 AA; 58174 MW; A7489BABEFA38F82 CRC64;

Query Match 3.1%; Score 135; DB 1; Length 501;
Best Local Similarity 20.9%; Pred. NO. 0.053;
Matches 98; Conservative 65; Mismatches 164; Indels 142; Gaps 26;

Oy 180 YEEYVVL--KNEVA--RANHVEDYG-----DYRGDYVNGVDGYYS 218
Db 103 YKEVILCSKAETLAMEAGKSDSFLSPYQLIEFNKREFTYV-----GYO-- 150
Oy 219 RGQIEDVEHTFEIKPILYENHLAYRAKLMNAPSYISPICLPAHLGLDWMGRFTNL 278
Db 151 -----EHPYALDLDFE--PGVTYKVL-----QLFAEL 177
Oy 279 YSLVPE-----GKPNIDVTDAWVDMAQRIEKAERFVSVGLPNTQGFWNS 331
Db 178 KEAIIPLVKQVYASGNP-----DTSFTKAPKPKKOKELSLYFQELGYDF--- 224
Oy 332 MLTPGVNOKAVCHPTAMDGLKGDPRILMCTKYTNDFLTA-----HHEMGILOYD--MAY 385
Db 225 ---DGRIDETV--HPFAFTLRGQDVRV--TTRYDEKDFRTALFGTTHGCHAIYEONIDE 278
Oy 386 AAQFPLLRNGANEHGHFA-----VGE-----IMLSAATPRHHTKISGLSPDF- 428
Db 279 ALSGTNLSDGASMSGHSQSLEFNTIGRKNKHWTPTYKKIQGASPVQFEDISL--DDEV 336
Oy 429 QEDNTEINFLKQA--LTIVGLTPFTYMLEKRWVFKGEIPRDMKKWEMKREIYG 486
Db 337 RAINSKRSFLRVEADELTLYPLHITIRIELEK--AIFSNESVEDLPSLMNKYDYDG 393
Oy 487 VVEPVPHDETCDPASPSEFHVSNDSYFIRYTRTLQFOFQALCOQ-----AAKHGEP 539
Db 394 I---PQPDADGIIQDVHAGDGFYFSPYALGTYVTAQLKQKMLDELDPFADLLENGER 450
Oy 540 HHCDDISNSTEAGOKLFNMLRLKSEFWTTLLENVYAKMMNVRPLNTEPEPLTWL 596
Db 451 HPIK-----QWLEKVIHIGKRRKPLDIKDATG--EELNVAYLIDY 490

RESULT 12
GIGB_BACSU STANDARD: PRT: 627 AA.
AC P39118;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

This SWISS-PROT entry is copyright
 between 1988 and 1990 by the
 Swiss Institute of Bioinformatics

643 FRSSVAYAMROYFLKKVKNOMTI.EGEEDVRYANI.KBRTCEVUTYU.A.D. 000

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; Z25795; CA81040.1; -
DR EMBL; AF008220; AAC00214.1; -
DR EMBL; Z29119; CARI5076.1; -.
DR PIR; S36624; S36624.
DR Subtilist; BG10907; g19B.
DR InterPro; IPR000461; Alpha-amylase.
DR InterPro; IPR004193; Isoamylase.N.
DR Pfam; PF00128; alpha-amylase.N.
DR Pfam; PF02922; isoamylase.N.1.
KW Glycogen biosynthesis; transferase; glycosyltransferase;
KM Complete proteome.
FT ACT_SITE 309 BY SIMILARITY.
FT ACT_SITE 352 BY SIMILARITY.
FT ACT_SITE 420 BY SIMILARITY.
SQ SEQUENCE 627 AA; 73665 MM; 64BA0553b6767BA CRC64;

Query Match      2.9%; Score 125; DB 1; Length 627;
Best Local Similarity 18.6%; Pred. No. 0.37;
Matches 131; Conservative 86; Mismatches 225; Indels 264; Gaps 37;

QY 49 NYNTITENVVNNNAAGDWSAFL-----KQSLLAQMYPLQLQNLTKVLQAQLQN 103
Db :::: |::|:: |::|:: |::|:: |::|:: |::|:: |::|:: |::|::
QY 56 DFNMSGEHHVHRVDNGCMTWTLFIFGIEKKR-----YYEIVTN-----N 97
Db :::: |::|:: |::|:: |::|:: |::|:: |::|:: |::|:: |::|::
QY 104 GSSVLSLDKSKRLNTLTMTSTYSTGKYCNDDNPQCCLLEPGLNEIMAN--SLDNER 161
Db 98 GEIRLKADP-----YAIYS-----EVRRNTASLYTDEE 125
QY 162 LMAMESRSVESEKOLRPLEYEYVVKENMARANHEDGGYWRGDYVNVGDIYSRQQ 221
Db :::: |::|:: |::|:: |::|:: |::|:: |::|:: |::|:: |::|::
QY 126 GYSWDQDKMQKQAKATLVKEKPEFL-----YEHLGSWK-----KHSDGRHSYAE 171
Db 122 LIDYD-----EHTEF--ELKPLYEHLAAVRAKIMANVPYSISPIGCLPAHLLGDWGAF 274
QY 172 LSOTLLPYTKKGTFHIELLPYE--HPYDRS-----WGQR 205
QY 275 WTNLVSLVPEFOQKNIDVTDAWDOA-----WDAQRFKKAKEKFVSVGLPNMF 324
Db 206 GTGYYSPTSRFG--PPHDLM-KFYDECHQONIGYIILDWVGPHCKAHGLYNFDGP-- 259
QY 325 QGVWEKSMILTDPGNQOKAVCHPTAMDLAGKDGFRTLMCTKTVMDDFLTAAH---HEMGHQ 380
Db 260 --LLEYKEERDEN-----WLGTANFDL--GREVNSFLTSALYWAFFYIH- 303
QY 381 YDMAVAAPPELLRNGANGFH-EAVGIMSLSAATRKLHLSIGLLSPDFOEDETETINFL 439
Db 304 -----DGERVDNAVANIILWPMDQERH-----TNPAVDLF 333
QY 440 LKQALTIVGLPFTYM--EKRWMMVFKEGIKPD-----OMMKKWMMEKKEIVGVBPV 491
Db 334 KLIINTREAREAPHYMIADSTEMPQYTGAVEGSGEGFIYKNNKGW--ANDVIKYMETP 390
QY 492 PHDETVCOPALEFHVNDYSFIRYTRTYLQOFQDALCOAKNHGRLPKDCDISNSTEAG 551
Db 391 PEERRHC-----HQLISFSL-----YAFSEHFVLPSHDEVYV-----G 425
QY 552 QKLENMLRLGSESEPWTALLENVGAKNMVVRPLLNV-----FEPLFTWLK 566
Db 426 KK--SLLNKMPGDY-----QKFAOYRIILLGTMVTHPGKLLIFWGSEFAQFDEM-K 473
QY 597 DQNKNSVVGSTWMSPADOST-----KVRLSLKSLGDKAYEWND--NEMYL 642
Db 474 DTGE--LDWFIDSFPMHOKASYFTODLLRFYQSKLILEYHDHAQSFEMVIDVHNDSOT 530
QY 643 FRSSVAYAMQYVLAKNQNMILFGEDRVANALKPRISNFETVAR 688
Db 531 F-SPIRGQKH-----GEALVITICNFPVVYHAYDVGP 563
```

```

RESULT 13
EP1B_STAEF
ID EP1B_STAEF STANDARD: PRT: 986 AA.
AC P310195;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Epidermin biosynthesis protein epib.
GN EP1B.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
NCB1_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RU 3298 / DSM 3095;
RX MEDLINE=9215237; PubMed=1740156;
RA Schnell N., Engelke G., Augustin J., Rosenstein R., Ungermann V.,
RA Goetz F., Entian K.-D.;
RT "Analysis of genes involved in the biosynthesis of lantibiotic
epidermin."
RL Eur. J. Biochem. 204:57-68(1992).
CC -1- FUNCTION: INVOLVED IN THE POSTTRANSLATIONAL MODIFICATION OF THE
ANTIBIOTIC EPIDERMIN.
CC -1- SUBCELLULAR LOCATION: POSSIBLY ASSOCIATED WITH, AND ANCHORED TO,
THE CYTOPLASMIC SIDE OF THE MEMBRANE.
CC -1- SIMILARITY: TO B.SUBTILIS SPAB AND L.LACTIS NISB.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

QY 384 AYA-----AOFELLRNGANGFEHANGEIMS----- 409
DB 644 IIAFGDNRLNLNDKHLILKLEKKHGRIRILESPINSENNRMEIYVPLYKKTSL 703
QY 410 --LSAATK-----HLK---STGLSPDDQEDNETENPLKALITVGTLPPTMYLE 457
DB 704 KQCSFTIIPKRNKRNKHNKDMFSLHSIPKTYQD---NFQDYLLPFTLEKVNPN 758
QY 458 KRWMYEKEIKKQDMKMKWEMKREIYGVPEVPHDEYCDPASLPHVSNDFIRY- 516
DB 759 KEFYIKFKED--EDFIK-LRLRE-----DEY-----SQIYSIRKWK 794
QY 517 -----TRITYOPQFQALCOAKHGPPLKCDISNSTAGOKLFNMLRGSEPTALE 571
DB 795 DYCLINSELYDSIYDYVEYRYGGPHYEDIEENFPWYDSL--SIMIQSE-FKIPKE 851
QY 572 NVYGAKNMNVRLNTEPELFTWLDQKNKSVGMSDMSPIADDSIVKRSKALDQ 631
DB 852 FIYA--ISIDFLDYLE-----INSEKEEILINNA--ED 882
QY 632 AYEMNDNEMYLERRSSVAVARQYFLKVNQMLFGEDYRVANLKPRISENFVYAPKNV 691
DB 883 LKRSND-----IREY-----KNLAKL-----INPKND 905
QY 692 SDIIP-----RTEVEKAIKRSRINDAF--RINDSLEFLGIQ 728
DB 906 YELKKEPFLNHEFLFNKISILENLKTLQKSLYSRSISGIFIMRCN--RIFGIN 961
QY 729 P 729
DB 962 P 962

```

```

RESULT 14
BGL_BACME
ID BGL_BACME STANDARD: PRT: 1034 AA.
AC 052847;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Beta-galactosidase (EC 3.2.1.23) (Lactase).
GN BGLM.
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCB1_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 319;
RA Strey J.;
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
galactose residues in beta-D-galactosides.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

DR PRINTS; PRO0132; GLHYDRLASE2.
DR PROSITE; PS00719; GLYCOSYL_HYDROL_F2_1; 1.
DR PROSITE; PS00608; GLYCOSYL_HYDROL_F2_2; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 481 481 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 547 547 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 1034 AA; 118673 MW; 38644C9A649415E9 CRC64;

Query Match
Best local similarity 19.2%; Pred. No. 1.4; Length 1034;
Matches 104; Conservative 66; Mismatches 165; Indels 206; Gaps 30;

QY 29 LKFNHEADLFYQSSL-----ASWNYNTITEENVQNMN---NAGDKSAFLKE 75
DB 260 LNFVFEHDQDLFEVVLVDANAQEVLPQNLNLSVSDQRTSLTHIKSPAKMSA----- 315
QY 76 QSTLQAMPYLOLQNL-TRYKLDLQALQNGSSVLSBDSK---RLNTLINTSTYISGK 131
DB 316 -----ESPMLTYLVL---SLKMAAGSIETESCKVGFRTFEIKNGIMLT--NGK 359
QY 132 -----VCPNPPOECILLEPGLNEMANSIDYNERLM----- 163
DB 360 RVLQGVNHERDYSKGRAGITREDMINDILMKQHINAVRTSHYPMDSVWYELCNEYG 419
QY 164 -----AW-----ESWSEVGKQLRPLYE----- 182
DB 420 LVYIDETNLEHGTWYTLQEGEOKAVPGSKPEMKENVLDRCRSMYERDKNHPSTIIISLG 479
QY 183 -----YVVLK-NEMARANHRYDYGDMVSGDYGVDCYDYSKQGLIEDVHT 229
DB 480 NESFGENFQHMHTFFEKDSTRLVNHYE--GIFFHRDSDASDIETMYVKP--ADVE-- 532
QY 230 FEELKPLYEHLHAYVRAKLMNAPSYISPIGCLPAHLIGDMWGRFTNLSYLVPEFGKP 289
DB 533 -----RYALMPKPKRYIL--CEVSHAMGNSCG--NLKYWELFDQYP 570
QY 290 NID---VTDAMVDQAMDQRIFFKKE--KEFVSVG-----LPNTQGFENSMKLTDPGNV 339
DB 571 ILQGFITWD-----WKDQALQATAEDETSTYLAVGSGDGPDPN--DGNCGGLIFADGTA 623
QY 340 QKAV-----CH-PTAW---DLGKGFRLMCTKYTMDFLTAHENGHI-----QYD MAY 385
DB 624 SPKIAEYKCYQPYKWTAVDPAKGFV-----QNKHLFTNLAAYDFVW 667
QY 386 AAQPEPLRNGANEGFHEAVGEIMSLSATPRHLKSGIGLSPDFQEDNETEINFLKQALT 445
DB 668 IVE---KNG-----ELVEKHASLNLNVA PDGDELFLSYLVYQENMETD- EFLVLTLSLR 716
QY 446 I 446
DB 717 L 717

RESULT 15
OSHS2_YEAST
ID OSH2_YEAST STANDARD; PRT; 1283 AA.
AC Q12451; P89891;
DF 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oxyesterol-binding protein homolog 2.
GN OSH2 OR YDLO19C OR D2845.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RA Ureastarazu L.A., Andre B., Vissers S.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SUBCELLULAR LOCATION.

```

```

RX MEDLINE=21301806; PubMed=11408574;
RA Levine T.P., Munro S.;
RT "Dual targeting of Oshp, a yeast homologue of oxyesterol-binding
RT protein, to both the Golgi and the nucleus-vacuole junction."
RL Mol. Biol. Cell 12:1633-1644(2001).
RN [3]
RP GENETIC ANALYSIS.
RX MEDLINE=21135676; PubMed=11238399;
RA Beh C.T., Cool L., Phillips J., Rine J.;
RT "Overlapping functions of the yeast oxyesterol-binding protein
RT homologues."
RL Genetics 157:1117-1140(2001).
CC -1 SUBCELLULAR LOCATION: Cytoplasmic. When bound to oxysterols it
CC translocate to the periphery of Golgi membranes.
CC -1 SIMILARITY: BELONGS TO THE OSBP FAMILY.
CC -1 SIMILARITY: CONTAINS 2 ANK REPEATS.
CC -1 SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
DR EMBL; Z74066; CA98577.1; -
DR EMBL; Z74067; CA98578.1; -
DR EMBL; Z48432; CA988340.1; -
DR SGD; S0002177; OSH2.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000648; Oxyesterol_BP.
DR InterPro; IPR001849; PH.
DR Pfam; PF00023; ank; 2.
DR Pfam; PF00169; PH; 1.
DR SMART; PF01237; Oxyesterol_BP; 1.
DR SMART; SM00248; ANK; 2.
DR SMART; SM00233; PH; 1.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 2.
DR PROSITE; PS01013; OSBP; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Lipid transport; transport; ANK repeat; Repeat; Golgi stack.
FT REPEAT 106 134 ANK 1.
FT REPEAT 206 235 ANK 2.
FT DOMAIN 289 386 PH.
SQ SEQUENCE 1283 AA; 145795 MW; D521957460E7F7C3 CRC64;

Query Match
Best local similarity 18.9%; Pred. No. 2.5; Length 1283;
Matches 171; Conservative 123; Mismatches 321; Indels 289; Gaps 43;

QY 10 SLVAVTAOSTIEBOAKTFLDKFNHEADLFYQSSLASW---NYNTITEENVQ---NMN 63
DB 412 ALISHSKTQSGSLPEASQYQHTLHKE---VIPSVSILYRRPSNMLSVSSSEIQLDNMLT 468
QY 64 MAGDKSAFLKQSTLAQYKYP-----LQEIQLYKLDLQALQNGSSVLSB-DKSK 114
DB 469 ESKGRVSAKMEINRLDGSKTGVTGSAALQVRSNMT-LKSNRMSQSGVSAIPIDKVP 527
QY 115 RLNTLINTSTYISY-----GKYCPNPPOECILLEPGLN-----EIMNS 155
DB 528 NCANLSQSNITTTGSAASLDNNYINDFEGDEANSDEEDL---GINFDEEYTKAQY 583
QY 156 LDYNERLAWESWRS-----EVGKQLRPLYEEVVLKNEAR----- 192
DB 584 GPYKKEKLDYEQALISIELSLIELIEQEPSEVWLTKKSLINTSTIFGLKDLTYKRD 643
QY 193 -----ANHYEDYGVWGDYEVNGVDYDSKQGLIEDVHTFEELKPLYEHLHAYVRA 246
DB 644 KRLVDVMSKQGVNVMW-----VQSVKLELMELSKTERLASIDKERGLKKIHL---K 694
QY 247 KLMNAPSYISPIGCL-----LPAILLDGMWGRFTNLSYLVPEFGKPNIDVTDAM 297

```

```

DB 695 KLESHATAGNCKSLNDEKQESDTASTGQI-----AKETISAT-----KEDESDA- 743
DB 298 VDOAMPDAORITFEAEKEFEVSVGLPNTQGWENSMITDPGNOKAVCHPTAMDIGKDR 357
DB 744 -DEFYDAELVDE-----VTELETAHPELSTAAKPHAPPV--PNEPD----- 784
DB 358 ILMCTVATMDFLTAHHEMCHIOYDMAVAAQPLPLNGANECHHEAVGELMISATPKH 417
DB 785 -----NDQYVDEKSIENVEKTSQKEKONNL-----VTEDEPKTDS 825
DB 418 LKSTIGLSPPOENETEIFLLKO-----ALITVGLPFTYMLEKRWMMVK 465
DB 826 LKMK-----AEDKESQVKEKTEKLAASSVIGEKTYAVTTVOKRKEEYLLKGGSTIGE 879
DB 466 GEIPKDOMMK-----WMEKREITG-----VPEPVPDETYCDPASLPHVNDYSFL 513
DB 880 DGIKRLSMKDKDPKISMAVILKSMVKGDMTWTLPV-----TFNEPTSLQ----- 927
DB 514 RYITRTLYQPOFALCOAAKHE-----GPIHCDSNSTEAGQKLFNMLRG--- 561
DB 928 -----RVAEDELEYSELDOAAATFEDSTLRTLYVAAFTASSVASTTKRYAKFPNPL-GEET 982
DB 562 -----KSEPWTLENNVYAK-----NMN 580
DB 983 EYSPDKQYRFTQVSHHPPISTATWTESPMDWGESEFVDYKNGSFVKKHGLMHK 1042
DB 581 VRPLNTYFEPFLTWLKDONK--NSFVGMSTWSPYADQSIKVRISLKSALGD-----KA 632
DB 1043 LRPDNEKEELTYWKKNFTVIGILG-----NPOVDNHEEVN--VNHGTGDHCKLYEKA 1096
DB 633 YENDNEMVYLFRRSSVAAAMR-QYFLKYNOMLLEGEEDVAVANKPRISTNEFTAPKNV 691
DB 1097 RGM-----RSSGAEITGEVYNNKKQKWLIGH-----WKEAIFAKKV 1136
DB 692 SDIIPREVEKAIKRSRIND-----AFRLNDSLEFLGIOP---TLGPPNPV 739
DB 1137 KD--GDSLEKT-RTASAGSGPDDGKFLIKWAKNRPREFPLTFPALTILNP-QPHL 1192
DB 740 SIWL 743
DB 1193 LPWL 1196

RESULT 16
YAMB_SCHPO STANDARD; PRT; 3655 AA.
ID YAMB_SCHPO STANDARD; PRT; 3655 AA.
AC Q10064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein CLF5.11c in chromosome I.
GN SPAC1F5.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Godle A., Hamlin N., Harris D., Hiderg J., Hodgson G.,
RA Holtby S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Mouton S., Mungall K., Murphy L., Muddett D., Odeli C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

```

```

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymopiez B.,
RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Filz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito G.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.,
RA "The genome sequence of Schizosaccharomyces pombe."
RA Nature 415:871-880(2002).
CC -1- SIMILARITY: STRONG, TO YEAST YHR099W.
CC -1- SIMILARITY: BELONGS TO THE P13/P14 KINASES FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: 268136; CAA92239.1; -.
DR InterPro: IPR003151; FAT.
DR InterPro: IPR003152; FATC.
DR InterPro: IPR000403; P13_P14_kinase.
DR InterPro: IPR001440; TPR.
DR Pfam: PF00454; P13_P14_kinase; 1.
DR Pfam: PF02259; FAT; 1.
DR Pfam: PF02260; FATC; 1.
DR SMART: SM00146; P13KC; 1.
DR PROSITE: PS0290; P13_4_KINASE_3; 1.
DR Hypothetical protein; Translase; Kinase.
FT DOMAIN 3324 3655 P13K/P14K (BY SIMILARITY).
FT SEQUENCE 3655 AA; 420774 MW; 50475E3F3FC2124A CRC64;

Query Match 2.8%; Score 119.5; DB 1; Length 3655;
Best Local Similarity 17.0%; Pred No. 11;
Matches 93; Conservative 93; Mismatches 197; Indels 165; Gaps 23;

DB 4 SSWLLSLVAVTAQSTIEBOAKT---FLDKFHEADELFY---OSSLASNNYNTITE 56
DB 2633 SDW-----STEDPLEKATKSLSPFSLRHRTADALLYLNKQKKGSVTEFSRIID 2684
DB 57 ENVOANNNAGDKASFLKE--OSTLAQMPLOLQNYLKYLOLQALQONGSSVLSDEKSK 114
DB 2685 ECKQFSLR--RWQOLPKRYQSHVSLHNFQELVEL-----QEAFGIYSQINDT 2731
DB 115 RLNTILNTMSTIYSTGKVCNDPDCILPEGLINEMANSIDYNERLMA--WESWSEVG 173
DB 2732 NIHHIDKL-----RDIKVVLOQWRELPVWD--DIDIMSDLTAMQSVF 2775
DB 174 KQLRLYEYVYLKNEKARNHEDYDGRGDEVGVGVYYSKQGLIEDEHFFEL 233
DB 2776 KSKINKVFLPLVSLAOSTNKSNTNSVYLXRGHEL----- 2811
DB 234 KPLEVHNAVYRAKIMAYPSYISPICLPAHLIGDMWGRFNTNLXSLVFPQOKPNIDV 293
DB 2812 -----AMINRF-AHARVHNHLEVCINOL-----IKIYTL-----PNIET 2846
DB 294 TDAMVDOAMPDAORITFEAEKEFEVSVGLPNT-----QGWENSMITDQG 337
DB 2847 QEATFLKREQEHCYESPSEQMLGLEVINNTNMTAFENRROKAEFTLKGFEON-RIGEKD 2905
DB 338 NVQAVCHPTAMDIGKDFRIMCTKYTMDFLTAHHEMCHIOYDMAVAAQPLPLNGAN 397
DB 2906 EANOAFATVAOIDIGSG-----KASEMG-LIHDELFOANPQEIHNACN 2948
DB 398 -EGFHEAVGEIMLSATPKHKLKSGISLSPFQEDNETELINLLKQALTIYGLPTYM 455

```

```

Db 2949 AVSCFLQA--SSLLSSNSKPLTRVILMLVSDSHGSEV-----YSSFKSE 2995
Qy 456 LKRMVAVKGEIPKDDMMKKMKRELVGVEPVHDEYCDPASFHVNDY-SFTR 514
Db 2996 IFPMWIKP---IPQ-----LLSALSHRESIHARAILIOIAKYPOSTLH 3036
Qy 515 YVTRTLQY 522
Db 3037 FOLRTAYE 3044

RESULT 17
DHC_PARTE STANDARD; PRT; 4540 AA.
ID DHC_PARTE STANDARD; PRT; 4540 AA.
AC Q27171;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Dyein heavy chain, cytosolic (DHC).
GN DHC-8.
OS Paramacium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillida;
OC Paramacium.
OX NCBI_TaxID=5688;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Stock 51;
RX MEDLINE=96157890; PubMed=8589455;
RA Kandi K.A., Forney J.D., Asai D.J.;
RT "The dynein genes of Paramacium tetraurelia: the structure and
RL expression of the ciliary beta and cytoplasmic heavy chains.";
RL Mol. Biol. Cell 6:1549-1562(1995).
CC -1- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A
CC MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND
CC ORGANELLES ALONG MICROTUBULES.
CC -1- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
CC INTERMEDIATE AND LIGHT CHAINS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U20449; AAA75445.1;
DR InterPro: IPR004273; Dynein_heavy.
DR Pfam: PF03028; Dynein_heavy; 1.
KW Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.
FT DOMAIN 440 482 COILED COIL (POTENTIAL).
FT DOMAIN 698 722 COILED COIL (POTENTIAL).
FT DOMAIN 794 827 COILED COIL (POTENTIAL).
FT DOMAIN 975 995 COILED COIL (POTENTIAL).
FT DOMAIN 1169 1251 COILED COIL (POTENTIAL).
FT DOMAIN 1295 1311 COILED COIL (POTENTIAL).
FT DOMAIN 1933 1961 MICROTUBULE-BINDING (POTENTIAL).
FT DOMAIN 3076 3182 COILED COIL (POTENTIAL).
FT DOMAIN 3289 3367 COILED COIL (POTENTIAL).
FT DOMAIN 3653 3688 COILED COIL (POTENTIAL).
FT DOMAIN 3820 3851 COILED COIL (POTENTIAL).
FT DOMAIN 4238 4259 COILED COIL (POTENTIAL).
FT DOMAIN 4313 4342 COILED COIL (POTENTIAL).
FT NP_BIND 1835 1842 ATP (POTENTIAL).
FT NP_BIND 2129 2136 ATP (POTENTIAL).
FT NP_BIND 2496 2503 ATP (POTENTIAL).
FT NP_BIND 2834 2841 ATP (POTENTIAL).
SQ SEQUENCE 4540 AA; 528626 MW; 0C6103148B8B95F9 CRC64;
Query Match 2.8%; Score 119.5; DB 1; Length 4540;

```

```

Best Local Similarity 18.3%; Pred. No. 15;
Matches 159; Conservative 131; Mismatches 315; Indels 265; Gaps 46;

Qy 12 VAVTAOSTTEEOAK-----TFLDENHEADLF-----YSSLAS 47
Db 521 IAISEVATITTNRDSLSGSAKEMFRLAFN---KLEFRPKAIGQYQSL-- 574
Qy 48 WNTYNTTEENVQNMNAGKWSAFLKEOST---LAQMPLOE---IQNTVQLQALQ 101
Db 575 ---LKTVHKDIQSLQKKE---TYQKQNSRLASARDIPLTSGFVMSQDIQRQKYM 628
Qy 102 QNGSSVL---SEPKSRILTIINTSTIYSTGKVCNPNPQCLLEGLNFIANSUD 157
Db 629 QKVEQIIGPQWAD-----TDEKKCKE-----METFERILD 660
Qy 158 YNRLMAMESRSEVG-----KQLPLVEEYVVLNEMARAHYEDGYRGGVEYNG 211
Db 661 SGPAL---EDWQKEINHNKAVSQNKLEPVYTRRGLEIRVYKESLSQLK---EVAN 714
Qy 212 VDG---YDYSQQLIEDVHEFEIKPLYEHLAYVR-----AKIMNAPSYISP 258
Db 715 LSNMKTQVPYSISHIANDAKASYFALSIOESLHYIQITSOLMAKSAKVALAKREVQ- 773
Qy 259 IGLPLHLLGDMWGRFTNLSLTVPEGQKPNIDVTDM--VDQMDA---QRIFK----- 309
Db 774 ---LQIQGFNYLTHTKTOLO-PYVK---FTDKVELEQAVNGINERIGQIESLC 822
Qy 310 EAEKFFVSVGLPMTQGFENSMITDPGVQKAVC-----HPYAMDGKGFRLMCT 362
Db 823 EAMKTCPPVSLD-----KLKDQEVDSICFNNSNLMHMQIDDK-QIESIICD 872
Qy 363 KVV---MDPL-----TAHMKGHQYDMMAYAP-----FLIRGAN 397
Db 873 RYTVQKEMLNQNTYQKIOERGLVNOQTVHEL-KLDDQIIVDPVEVAKYFWQOE-- 928
Qy 398 EGFHEAVGEIMSLATPKHL-----KSIGLSPDQEDNETEIN---FLIKQALTVG 448
Db 929 -FHKMIGICSLPRVYANRFNTIQNGPMTGROLDSTTIKINQOLKDAYSOIG 986
Qy 449 TLPTFYMLEKRWVFKGIPKDDMMKW---WEKRELVGVEPVHDE----- 495
Db 987 QL---LE-----DMEQYQVQWLNQYSLMEID---IKQVEDLDDIEKQWQM 1027
Qy 496 ---TYDDPASLHF---VSNDSIRIKYTTITLYQFOQELCAAHHEGRLKCD 543
Db 1028 LVDIKQGRATFNSSTEEHFGAIIIDYRVQYKINHXYDAMHKKEL----- 1073
Qy 544 ISNSTEAGKLFNMLRGLKSEFWTLALENVYAK---NMNVRPLNVEEPLFTWLKDN 599
Db 1074 ---NHFGNKFGEQLRVFN-----KNVTEKELKLKINFQDLTSIISITIIQED 1121
Qy 600 KNSPFGWSTDSNPYAD-OSIKVRLSK---SALGKAYEWNDENMYLFRSSVAYAMROYF 655
Db 1122 K-KFGWMSADIESFNGQKVDRORYQYQDMLSEGEVQWQNFQKRLSKKLSQSESM 1180
Qy 656 LKVKQMLIFGEDEVYVNLKPRISFNFEVYAPKVVSLIIR---TEVEKARMSRIND 713
Db 1181 NNIOQKI---QODRYLNOQIQLIEQWKTSKPQSGCSPNEAEDILKSLNQLISVQE 1236
Qy 714 AFLRLNDSLLEPLGIQTPGPNQPVSYIWL 743
Db 1237 KYECSQAKKELIKMP---PTHQOKLVNLL 1263

RESULT 18
MTHR_CAEEL STANDARD; PRT; 663 AA.
ID MTHR_CAEEL STANDARD; PRT; 663 AA.
AC Q17693;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable methyltetrahydrofolate reductase (EC 1.5.1.20).
GN C06A8.1.

```

OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPRAIN-Bristol N2;
 RA Leimbach D.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADPH.
 CC methyltetrahydrofolate + NADPH.
 CC -1 COFACTOR: FAD (By Similarity).
 CC -1 PATHWAY: Folate metabolism.
 CC -1 SIMILARITY: BELONGS TO THE METHYLENETERAHYDROFOLATE REDUCTASE
 CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U39849; AAA81048.2; -
 DR HSSP: P00394; 1BS7.
 DR WormRep: C06A8.1; CE30593.
 DR InterPro: IPR004621; Fadh2_euk.
 DR InterPro: IPR003171; MethylotF_redctse.
 DR Pfam: PF02219; MTHFR_1.
 DR TIGRFAMs: TIGR00677; fadh2_euk; 1.
 DR Oxidoreductase; Flavoprotein; FAD; NADP.
 KW SEQUENCE 663 AA; 75486 MW; EA94806A2C3BC1CD CRC64;
 SQ
 Query Match 2.7%; Score 118; DB 1; Length 663;
 Best Local Similarity 18.3%; Pred. No. 1.2; Indels 210; Gaps 30;
 Matches 101; Conservative 77; Mismatches 163;
 QY 305 ORIFKAERF-----VSGLPN-----MTGG-----FWNSMLTPGNNOK- 341
 DB 63 ERLIDKQPFSLFFPRFVNGVPRFLERLVERLSGSGSVFVDMWH--MGSDPAVNVKV 120
 QY 342 -----AVCHPTAMDLGKD----- 355
 DB 121 TSSSTIAASMLDYCGVDTMLHMTQVYNKADTLKHLDEAKKAMGLRSIALRGDLPPTTEL 180
 QY 356 -----FRILMCTFVTMDDEL-----AHHEMGHIQYDAVAAQAPFLRNKANGFHEAV 404
 DB 181 EDTHQFRALDMIRWIREEGNYSIGCAGYPLGHQAP--SYKADMLYKAKCDAGANFYI 239
 QY 405 GEIMLSAATPKHL--KSIGLSPDFQEDNETEINFLKQALITVGLPT--YMLEKRR 460
 DB 240 TQLEFAEFERKEFVROCRELIGTP-----IIPGIMPIMGESIKRI 281
 QY 461 WNFYKGEIPKQWKKMKKEKREIVGVVEPRPHDETICDPAFLHNSNYSFIRXYTRRL 520
 DB 282 AKLSOLEIP-----OHLLDLLEIRKD-----DDAVOKYGTERC 315
 QY 521 YQFOFOALCOAKHGEPLKHCDSNSTEAGKLFNLRLGKSEP-----WTLLENV 573
 DB 316 IE-----MCRRL-----LDNQTASHILHYTNRRGSLIREILKSLGLMKLGGDRY 359
 QY 574 VGAKNMVRLP--LNFEPFLFWLKDONKNSFVGWSTDSFVYDQSLKVAISLSKALGDK 631
 DB 360 FPMKNSQHPIRLCLESVRLY-W--SFRPRSITRTDMQDFPNG----- 401
 QY 632 AYENDNEMYLFRSSVAYV--MROVELK-----VKN--QMLFGF-----EDVRVANKLPR 678
 DB 402 --RWGNS-----SSPAFGDVSSYILSNLTITVRNADRLAMFGANIESFEDVK-----R 447

QY 679 ISNFVETAPKNSVD-----IIPREVEKAIMSSGRINDAF-RINDNSLEFLGIQPTL- 731
 DB 448 VFVNIYITQAPN--ADGVKVTLEMTAEATGVQPEFSLISEQLWYCNENIILVNQSPSVN 505
 QY 732 GPNQPPVSIW 742
 DB 506 GAPSTDPDVGW 516
 RESULT 19
 ID AKA9_HUMAN STANDARD; PRT; 3911 AA.
 AC 099996; 09U004; 09U0H3; 09Y6Y2; 014869; 043355; 094895; 09Y6B8;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE A-kinase anchor protein 9 (protein kinase A anchoring protein 9)
 DE (PKA9) (A-kinase anchor protein 450 kDa) (AKAP 450) (A-kinase anchor
 DE protein 350 kDa) (AKAP 350) (hAKAP 350) (AKAP 120 like protein)
 DE (Hyperin protein) (Yotiao protein) (Centrosome- and golgi-1-located
 DE PKA-associated protein) (CG-NAP).
 GN AKAP9 OR AKAP450 OR AKAP350 OR KIA0803.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 4).
 RC TISSUE=Brain;
 RA Loh J.W., Wyszynski M., Madhavan R., Sealock R., Kim J.U., Sheng M.;
 RT "Yotiao, a novel protein of neuromuscular junction and brain that
 RT interacts with specific splice variants of NMDA receptor subunit
 RT NR1".
 RL J. Neurosci. 18:2017-2027(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND VARIANT GLN-1347 INS.
 RC MEDLINE=9219864; PubMed=10202149;
 RA Witczak O., Skaalheg B.S., Keryer G., Bornens M., Tasken K.,
 RA Jansen T., Gerstlitz S.;
 RT "Cloning and characterization of a cDNA encoding an A-kinase anchoring
 RT protein located in the centrosome, AKAP450".
 RL EMBO J. 18:1858-1868(1999).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC TISSUE=Brain;
 RA Takahashi M., Shibata H., Shimakawa M., Miyamoto M., Mukai H., Ono Y.;
 RT "Characterization of a novel giant scaffolding protein, CG-NAP, that
 RT anchors multiple signaling enzymes to centrosome and the golgi
 RT apparatus".
 RL J. Biol. Chem. 274:17267-17274(1999).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Kemmer W.A., Deliss S., Schwarz U.;
 RT "Cloning of Hyperion".
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 323-3911 FROM N.A. (ISOFORM 2).
 RC TISSUE=Gastric parietal cell;
 RX MEDLINE=99115654; PubMed=9915845;
 RA Schmidt P.H., Dransfield D.T., Claudio J.O., Hawley R.G.,
 RA Trotter K.W., Milgram S.L., Golderring J.R.;
 RT "AKAP350, a multiply spliced protein kinase A-anchoring protein
 RT associated with centrosomes".
 RL J. Biol. Chem. 274:3055-3066(1999).
 RN [6]
 RP SEQUENCE OF 1802-3876 FROM N.A. (ISOFORM 5).
 RC TISSUE=Lymphoid;
 RA Hinds R., Suterer C., Becker M., Hawkins M.;
 RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [7]

FT	DOMAIN	1585	1659	COILED COIL (POTENTIAL).
FT	DOMAIN	1857	2455	COILED COIL (POTENTIAL).
FT	DOMAIN	2544	2561	COILED COIL (POTENTIAL).
FT	DOMAIN	2603	2776	COILED COIL (POTENTIAL).
FT	DOMAIN	3065	3092	COILED COIL (POTENTIAL).
FT	DOMAIN	3124	3470	COILED COIL (POTENTIAL).
FT	DOMAIN	3587	3689	COILED COIL (POTENTIAL).
FT	DOMAIN	3726	3730	COILED COIL (POTENTIAL).
FT	DOMAIN	203	292	POLY-LEU.
FT	DOMAIN	321	1010	GLN-RICH.
FT	DOMAIN	1846	2772	GLU-RICH.
FT	VARSPPLIC	17	28	MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT	VARSPPLIC	1637	1642	OLOEPI -> LARRD (IN ISOFORM 4).
FT	VARSPPLIC	1633	3911	MISSING (IN ISOFORM 4).
FT	VARSPPLIC	2175	2182	MISSING (IN ISOFORM 3).
FT	VARSPPLIC	2175	2183	SADPQKVE -> Q (IN ISOFORM 6).
FT	VARSPPLIC	2895	2907	VAFGYNMFCTLC -> GSSIPLAHSDAYREICSS (IN ISOFORM 2, ISOFORM 3 AND ISOFORM 6).
FT	VARSPPLIC	2895	2948	MISSING (IN ISOFORM 5).
FT	VARSPPLIC	3901	3911	STOFHNGMR -> ALSLTTSWQHSHARPAPLPEELLISH SLG (IN ISOFORM 6).
FT	VARANT	1347	1347	K -> KQ.
FT	CONFLICT	76	76	/FtId=VAR_-010926.
FT	CONFLICT	475	475	E -> Q (IN REF. 3).
FT	CONFLICT	554	554	M -> I (IN REF. 3).
FT	CONFLICT	638	638	E -> G (IN REF. 3).
FT	CONFLICT	663	663	R -> S (IN REF. 3).
FT	CONFLICT	913	913	N -> S (IN REF. 3).
FT	CONFLICT	956	956	H -> N (IN REF. 3).
FT	CONFLICT	980	982	K -> N (IN REF. 3).
FT	CONFLICT	997	997	OKH -> PKR (IN REF. 1 AND 2).
FT	CONFLICT	1001	1001	Q -> P (IN REF. 1 AND 2).
FT	CONFLICT	1020	1020	Q -> P (IN REF. 1 AND 2).
FT	CONFLICT	1028	1028	N -> D (IN REF. 3).
FT	CONFLICT	1626	1626	V -> E (IN REF. 3).
FT	CONFLICT	1703	1703	R -> P (IN REF. 1 AND 2).
FT	CONFLICT	1707	1707	N -> T (IN REF. 3).
FT	CONFLICT	1802	1803	V -> G (IN REF. 3).
FT	CONFLICT	1843	1843	MISSING (IN REF. 5).
FT	CONFLICT	1956	1956	A -> P (IN REF. 3).
FT	CONFLICT	2027	2027	I -> V (IN REF. 3).
FT	CONFLICT	2157	2158	V -> D (IN REF. 5).
FT	CONFLICT	2169	2169	EI -> HE (IN REF. 7).
FT	CONFLICT	2514	2514	E -> V (IN REF. 3).
FT	CONFLICT	2851	2851	L -> R (IN REF. 3).
FT	CONFLICT	2957	2957	E -> N (IN REF. 8).
FT	CONFLICT	2983	2983	E -> D (IN REF. 3).
FT	CONFLICT	3087	3087	P -> S (IN REF. 3, 7 AND 8).
FT	CONFLICT	3218	3218	Q -> H (IN REF. 3).
FT	CONFLICT	3307	3309	Q -> H (IN REF. 3).
FT	CONFLICT	3751	3751	ESE -> OSO (IN REF. 3).
FT	CONFLICT	3833	3833	P -> A (IN REF. 3).
FT	CONFLICT	3911	453664	T -> S (IN REF. 3).
FT	SEQUENCE	3911	453664	MM; 3FB1CB1C819B47AA CRC64;
QY	Query Match	2.7%;	Score 118;	DB 1; Length 3911;
QY	Best Local Similarity	20.2%;	Pred. No. 15;	
QY	Matches 70;	Conservative 63;	Mismatches 130;	Indels 84; Gaps 15
QY	13	AVTAAGSTIEQKTFDKFENHAEEDLFFYQSS-----LASNNYNINITEENVQANNAGDK	68	
QY	1095	SLRATQPSENDKQKRELNVKSEQNDLRLKQMAQRICISLVSITHVDVR-EYMEIEKKK	1153	
QY	69	WSAFKEQSTLAQMPLEQIONL-TVKIQLOALAQNG-----SSVLSPEKSK	114	
QY	1154	ALCSIEKEELFAQEEKIKELQKHQLELQTKQTDGDEGKPLHLGLKQKAVSEBCSY	1213	
QY	115	RLMTIINTMSTIYSKVC-----RPDNPQECLL--LEPGINETMANSIDYNE-----	160	
QY	1214	FLQDTLVSIGEYTPALKCEVNAEDKNSGYSISENEDPELDQRYEVQDFQENMHTLLN	1273	
QY	161	-----RLMAWES--WSEVSGKQLRLPYE-EYVVLKNEKARANHEDYG	200	

Db 1274 KYTEENKLVLTQRLSKIMGQTDGKMLFEGEENLPKETEFLSIHSGMNTLE----- 1327
 QY 201 DYMWDYEVNGVDYDYSRGOLIEVEHT-FBEIKPLYEHLHAYRAKLNAYPSYSPI 259
 Db 1328 -----DIVN-----HKSLSLIDLEKTKLEEOVELESLSIOOOLKETEONVEAEI 1377
 QY 260 GCLPAHLGDMWGRFMTNLYSLTFVFGQKPNIDVTDAMVDAMPAR 306
 Db 1378 HCLQRLQA-----VSESTVP-----PSLPDVSIVITES-DAOR 1410

RESULT 20
 UTRQ_HUMAN STANDARD; PRT; 3433 AA.
 ID UTRQ_HUMAN STANDARD; PRT; 3433 AA.
 AC P46939;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Utrrophin (Dystrophin-related protein 1) (DRP1) (DRP).
 GN UTRN OR DMDL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93096045; PubMed=1461283;
 RA Tinsley J.M., Blake D.J., Roche A., Fairbrother U., Riss J.,
 RA Byth B.C., Knight A.E., Kendrick-Jones J., Suthers G.K., Love D.R.,
 RA Edwards Y.H., Davies K.E.;
 RT "Primary structure of dystrophin-related protein.";
 RL Nature 360:591-593(1992).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 144-261.
 RA MEDLINE=99141377; PubMed=9887274;
 RA Keep N.H., Norwood F.L.M., Moores C.A., Winder S.J.,
 RA Kendrick-Jones J.;
 RT "The 2.0-A structure of the second calponin homology domain from the
 RT actin-binding region of the dystrophin homologue utrophin.";
 RL J. Mol. Biol. 285:1257-1264(1999).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 31-256.
 RA MEDLINE=20113481; PubMed=10647184;
 RA Keep N.H., Winder S.J., Moores C.A., Walke S., Norwood F.L.M.,
 RA Kendrick-Jones J.;
 RT "Crystal structure of the actin-binding region of utrophin reveals a
 RT head-to-tail dimer.";
 RL Structure 7:1539-1546(1999).
 CC -1- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE
 CC PLASMA MEMBRANE (BY SIMILARITY TO DYSTROPHIN).
 CC -1- SUBCELLULAR LOCATION: NEUROMUSCULAR JUNCTION.
 CC -1- TISSUE SPECIFICITY: MUSCLE.
 CC -1- SIMILARITY: STRONG, TO DYSTROPHIN.
 CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
 CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,
 CC ABP-120, ABP-180, OR BETA-FODRIN).
 CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 20 SPECTRIN REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 ZK-TYPE ZINC FINGER.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X69086; CAA4829.1; -
 CC PIR: S28381; S28381
 CC PDB: 1BHD; 16-FEB-99.
 CC PDB: 1QAG; 01-JAN-00.

DR Genew; HSCN:12635; UTRN.
 DR MTM; 128240; -
 DR InterPro; IPR001589; Actbind_actnln.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR002017; Spectrin.
 DR InterPro; IPR003122; Tatr.
 DR InterPro; IPR002349; WW.
 DR InterPro; IPR001202; WW_Rsp5_WWP.
 DR InterPro; IPR000433; Znf_ZZ.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00435; WW; 1.
 DR Pfam; PF00559; ZF_ZZ_1; 1.
 DR PRINTS; PR00403; WMDOMAIN.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00150; SPEC; 18.
 DR SMART; SM00319; TATR; 1.
 DR SMART; SM00456; WW; 1.
 DR SMART; SM00291; ZNF_ZZ_1; 1.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.
 DR PROSITE; PS01159; WW_DOMAIN_1; 1.
 DR PROSITE; PS50030; WW_DOMAIN_2; 1.
 DR PROSITE; PS01357; ZF_ZZ_1; 1.
 DR PROSITE; PS0135; ZF_ZZ_2; 1.
 DR PROSITE; PS0135; ZF_ZZ_2; 1.
 KW Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;
 KW Repeat; 3D-structure; Zinc-finger.
 FT DOMAIN 1 246
 FT DOMAIN 2 135
 FT DOMAIN 3 150
 FT DOMAIN 4 252
 FT DOMAIN 5 308
 FT DOMAIN 6 309
 FT DOMAIN 7 417
 FT DOMAIN 8 526
 FT DOMAIN 9 541
 FT DOMAIN 10 637
 FT DOMAIN 11 798
 FT DOMAIN 12 803
 FT DOMAIN 13 902
 FT DOMAIN 14 1016
 FT DOMAIN 15 1083
 FT DOMAIN 16 1125
 FT DOMAIN 17 1230
 FT DOMAIN 18 1248
 FT DOMAIN 19 1432
 FT DOMAIN 20 1541
 FT DOMAIN 21 1544
 FT DOMAIN 22 1652
 FT DOMAIN 23 1753
 FT DOMAIN 24 1910
 FT DOMAIN 25 1968
 FT DOMAIN 26 2081
 FT DOMAIN 27 2258
 FT DOMAIN 28 2333
 FT DOMAIN 29 2399
 FT DOMAIN 30 2440
 FT DOMAIN 31 2556
 FT DOMAIN 32 2636
 FT DOMAIN 33 2658
 FT DOMAIN 34 2688
 FT DOMAIN 35 2691
 FT DOMAIN 36 2797
 FT DOMAIN 37 2845
 FT DOMAIN 38 3064
 FT DOMAIN 39 3111
 FT DOMAIN 40 394488
 FT ZN_FING 3433 AA; EAE8DB409F858E5B CRC64;
 SO SEQUENCE

Query Match 2.7%; Score 115.5; DB 1; Length 3433;
 Best Local Similarity 18.5%; Pred. No. 19;
 Matches 139; Conservative 98; Mismatches 249; Indels 265; Gaps 36;

QY 29 LDFEHEAEEL-----FYOSSLASNNYNTNTEENVQMMNNAAGDKWSAFLKEQSTL 79
 Db 2533 LDMNORWMDLAKKSSSTRAHLEASAEKKNRLMSLEELIKWLNKDE---LKKQMTI 2588
 QY 80 AOMYPLQEIQNLTVKQLQALQONGSSVLEDESKRLNTILNTMSTIYSTGKVCNPDNQ 139
 Db 2589 GGDVPAIQOYDCHKALRELEKEKESVLANVDAQRY-----FLADQIEAPEPR 2639
 QY 140 ECLLEPGL-----NETMANSLDYNEKLAMES-----WRSEVGQLRPLYEE 182
 Db 2640 RNLQSTTELTPEERACKIAKARKOSSEYKER--WESLNVAVTSWQOYKALKEKLDL 2696
 QY 183 YVVLKNEMARANHYEDYGDYMR--GDYEYNGVDGYDSGOLIEVEHTFEIRKPLYEHL 240

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 16:53:52 ; Search time 36 Seconds
(without alignments)
4607.442 Million cell updates/sec

Title: US-09-978-385-2
Perfect score: 4291
Sequence: 1 MSSSSWLLSLVAVTAQST.....ISKGNNGFQNTDVGTSF 805

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPRENBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4291	100.0	805	4 Q9NRA7	Q9nra7 homo sapien
2	4288	99.9	805	4 Q9BYF1	Q9byf1 homo sapien
3	4281	99.8	804	4 Q9UFZ6	Q9ufz6 homo sapien
4	3579	83.4	805	11 Q8R010	Q8r010 mus musculu
5	3485	81.2	798	11 Q99N71	Q99n71 mus musculu
6	1539	35.9	353	11 Q99N70	Q99n70 mus musculu
7	1336	31.1	732	6 Q9GLN6	Q9glng pan troglod
8	1329	31.0	1304	6 Q9GLN7	Q9glng pan troglod
9	1310	30.5	1313	11 Q9EOM9	Q9eqm9 rattus norv
10	1303	30.4	694	4 Q15540	Q15540 homo sapien
11	1089.5	25.4	648	5 Q9NDS8	Q9nds8 hombyx mori
12	1086	25.3	615	5 Q9NKE4	Q9nke4 drosophila
13	1057	24.6	660	5 Q17248	Q17248 boophilus m
14	1030	24.0	630	5 Q24222	Q24222 drosophila
15	1028	24.0	630	5 Q9VLJ6	Q9vlj6 drosophila
16	1004	23.4	249	11 Q9D836	Q9d836 mus musculu

17	930	21.7	844	5 Q8SXX2	Q8sxx2 drosophila
18	841	19.6	792	5 Q9VJY1	Q9vjy1 drosophila
19	642.5	15.0	907	5 Q18581	Q18581 caenorhabdi
20	502.5	11.7	661	5 Q9V520	Q9v520 drosophila
21	476	11.1	611	5 Q9VJV2	Q9vjv2 drosophila
22	434.5	10.1	202	11 Q64603	Q64603 rattus norv
23	409.5	9.5	628	5 Q9W021	Q9w021 drosophila
24	408.5	9.5	628	5 Q95U59	Q95u59 drosophila
25	390	9.1	222	11 Q9ESG3	Q9esg3 rattus norv
26	387	9.0	222	4 Q9HBJ8	Q9hbj8 homo sapien
27	386	9.0	222	11 Q9ESG4	Q9esg4 mus musculu
28	306.5	7.1	121	6 Q9BDG1	Q9bdg1 bos taurus
29	280.5	6.5	135	4 Q16425	Q16425 homo sapien
30	251.5	5.9	157	11 Q61265	Q61265 mus musculu
31	239	5.6	775	5 Q9TX66	Q9tx66 drosophila
32	157	3.7	532	16 Q9KRV0	Q9krv0 bacillus ha
33	154	3.6	502	16 Q8Y616	Q8y616 listeria mo
34	152.5	3.6	734	5 Q9NKE3	Q9nke3 drosophila
35	147	3.4	54	6 Q95161	Q95161 canis famli
36	147	3.4	502	16 Q92AC3	Q92ac3 listeria in
37	139.5	3.3	987	16 Q8YWG6	Q8ywg6 anabaena sp
38	139	3.2	608	16 Q9PR80	Q9pr80 ureaplasma
39	136	3.2	611	16 Q9PPW8	Q9ppw8 ureaplasma
40	130.5	3.0	461	2 Q9K152	Q9k152 enterococu
41	127.5	3.0	3571	10 Q9SL27	Q9sl27 arabidopsis
42	127.5	3.0	3574	10 Q9AUB4	Q9aub4 arabidopsis
43	127	3.0	779	5 Q9XYS8	Q9xys8 dictyosteli
44	125.5	2.9	600	16 Q8R670	Q8r670 fusbacteri
45	124.5	2.9	987	11 Q61636	Q61636 mus musculu

ALIGNMENTS

RESULT 1
Q9NRA7 PRELIMINARY: PRT: 805 AA.
ID Q9NRA7
AC Q9NRA7:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Angiotensin converting enzyme-like protein (ACE-related
DE carboxypeptidase ACE2).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOMA;
RA Tipids S.R., Hooper N.M., Hyde R.J., Christie G., Karran E.,
RA Turner A.J.;
RT "A Human Homolog of Angiotensin Converting Enzyme - Cloning and
RT Functional Expression As A Captopril-Insensitive Carboxypeptidase.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Donoghue M., Hsieh F., Baronas E., Godbout K., Gosselin M.,
RA Stagliano N., Donovan M., Woolf B., Robison K., Jeyaseelan R.,
RA Breitbart R.E., Acton S.;
RT "A novel ACE-related carboxypeptidase (ACE2) converts angiotensin I to
RT angiotensin-9." (2000).
RL Cite. Res. 0:0-0(2000).
DR EMBL: AF241254; AAF78220.1; -;
DR EMBL: AF291820; AAF9721.1; -;
DR MEROPS: M02.006; -;
DR InterPro: IPR001548; Zn_MTPeptide.
DR InterPro: IPR001330; Zn_MTPeptide.
DR Pfam: PF01401; Peptidase_M2; 1.
DR PRINTS: PR00791; PEPTIDASE.
DR PRODOM: PD004184; Peptidase_M2; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Carboxypeptidase.

SEQ SEQUENCE 805 AA; 92462 MW; 8EE6EB0A931550E8 CRC64;

Query Match 100.0%; Score 4291; DB 4; Length 805;
Best Local Similarity 100.0%; Pred. No. 3.5e-305;
Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 M S S S W L L S L V A V T A A S T I E Q A K T F L D K F N H E A E D L F Y O S S L A S M N Y N T N I T E E N V Q 60
   1 M S S S W L L S L V A V T A A S T I E Q A K T F L D K F N H E A E D L F Y O S S L A S M N Y N T N I T E E N V Q 60
Db 1 M S S S W L L S L V A V T A A S T I E Q A K T F L D K F N H E A E D L F Y O S S L A S M N Y N T N I T E E N V Q 60
QY 61 M N N A G D K W S A F L K E Q S T L A Q M Y P L O E I Q N L T V K L O L A L O Q S S V L S D K S K R L N T I L 120
   61 M N N A G D K W S A F L K E Q S T L A Q M Y P L O E I Q N L T V K L O L A L O Q S S V L S D K S K R L N T I L 120
Db 61 M N N A G D K W S A F L K E Q S T L A Q M Y P L O E I Q N L T V K L O L A L O Q S S V L S D K S K R L N T I L 120
QY 121 N T M S T I Y S T G K V C N P D N P O E C L L E P G L N E I M A N S L D Y N E R L A M E S W S E V G K O L R P L Y 180
   121 N T M S T I Y S T G K V C N P D N P O E C L L E P G L N E I M A N S L D Y N E R L A M E S W S E V G K O L R P L Y 180
Db 121 N T M S T I Y S T G K V C N P D N P O E C L L E P G L N E I M A N S L D Y N E R L A M E S W S E V G K O L R P L Y 180
QY 181 E E Y V L K N E M A R A N H Y E D Y G D Y W R G D Y E V N G V D Y S R G O L I E D V E H T F E E I K P L Y E H L 240
   181 E E Y V L K N E M A R A N H Y E D Y G D Y W R G D Y E V N G V D Y S R G O L I E D V E H T F E E I K P L Y E H L 240
Db 181 E E Y V L K N E M A R A N H Y E D Y G D Y W R G D Y E V N G V D Y S R G O L I E D V E H T F E E I K P L Y E H L 240
QY 241 H A Y V R A K L M N A P S Y I S P I G C L P A H L D M G R F W T N L Y S L T V P G Q K P N I D Y T D A M V Q 300
   241 H A Y V R A K L M N A P S Y I S P I G C L P A H L D M G R F W T N L Y S L T V P G Q K P N I D Y T D A M V Q 300
Db 241 H A Y V R A K L M N A P S Y I S P I G C L P A H L D M G R F W T N L Y S L T V P G Q K P N I D Y T D A M V Q 300
QY 301 A M D A O R I F E A K E F F V S V G L P N M T O G F W E N S M L T D P G N V O K A V C H P T A M D L G K G D F R I L M 360
   301 A M D A O R I F E A K E F F V S V G L P N M T O G F W E N S M L T D P G N V O K A V C H P T A M D L G K G D F R I L M 360
Db 301 A M D A O R I F E A K E F F V S V G L P N M T O G F W E N S M L T D P G N V O K A V C H P T A M D L G K G D F R I L M 360
QY 361 C T K V T M D E F L T A H N E M G H I O Y D M A Y A A O P F L L R N G A N E G F H E A V E I N S L S A A P R K H L S 420
   361 C T K V T M D E F L T A H N E M G H I O Y D M A Y A A O P F L L R N G A N E G F H E A V E I N S L S A A P R K H L S 420
Db 361 C T K V T M D E F L T A H N E M G H I O Y D M A Y A A O P F L L R N G A N E G F H E A V E I N S L S A A P R K H L S 420
QY 421 I G L S P D F O E D N E T E I N F L K Q A L T I V G L P T Y M L E K W M V F E G E I P R D O M K K W M E M 480
   421 I G L S P D F O E D N E T E I N F L K Q A L T I V G L P T Y M L E K W M V F E G E I P R D O M K K W M E M 480
Db 421 I G L S P D F O E D N E T E I N F L K Q A L T I V G L P T Y M L E K W M V F E G E I P R D O M K K W M E M 480
QY 481 K R E I V G V E P V P H D E T Y C D P A S L F H V S N D Y S F I R Y T R T L Y O F O Q A L C O A K H E G P L H 540
   481 K R E I V G V E P V P H D E T Y C D P A S L F H V S N D Y S F I R Y T R T L Y O F O Q A L C O A K H E G P L H 540
Db 481 K R E I V G V E P V P H D E T Y C D P A S L F H V S N D Y S F I R Y T R T L Y O F O Q A L C O A K H E G P L H 540
QY 541 K C D I S N S T E A G Q K L F N M L R L G K S E P W T L A L E N V G A K N M N V R P L I N Y E P L E T W L K D O N K 600
   541 K C D I S N S T E A G Q K L F N M L R L G K S E P W T L A L E N V G A K N M N V R P L I N Y E P L E T W L K D O N K 600
Db 541 K C D I S N S T E A G Q K L F N M L R L G K S E P W T L A L E N V G A K N M N V R P L I N Y E P L E T W L K D O N K 600
QY 601 N S F V G M S T D M S P Y A D O S I K V R I S L K S A L G D K A Y E M N D N E M Y L F R S S V A Y A M R Q Y L K A V N 660
   601 N S F V G M S T D M S P Y A D O S I K V R I S L K S A L G D K A Y E M N D N E M Y L F R S S V A Y A M R Q Y L K A V N 660
Db 601 N S F V G M S T D M S P Y A D O S I K V R I S L K S A L G D K A Y E M N D N E M Y L F R S S V A Y A M R Q Y L K A V N 660
QY 661 Q M I L F G E E D V R A N L K P R I S F N F V T A P K N V S D I I P R T E V E K A I R M S R S R I N D A F R L N D N 720
   661 Q M I L F G E E D V R A N L K P R I S F N F V T A P K N V S D I I P R T E V E K A I R M S R S R I N D A F R L N D N 720
Db 661 Q M I L F G E E D V R A N L K P R I S F N F V T A P K N V S D I I P R T E V E K A I R M S R S R I N D A F R L N D N 720
QY 721 S L E F I G I O P T L G P P N O P V S I W L I V G V M G Y I V G I I L F T G I R D K K K K A S G E N P 780
   721 S L E F I G I O P T L G P P N O P V S I W L I V G V M G Y I V G I I L F T G I R D K K K K A S G E N P 780
Db 721 S L E F I G I O P T L G P P N O P V S I W L I V G V M G Y I V G I I L F T G I R D K K K K A S G E N P 780
QY 781 Y A S I D I S K G E N N P G F O N T D D V O T S F 805
   781 Y A S I D I S K G E N N P G F O N T D D V O T S F 805
Db 781 Y A S I D I S K G E N N P G F O N T D D V O T S F 805

```

RESULT 2

Q9BYF1 PRELIMINARY; PRF; 805 AA.
AC Q9BYF1;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE ACB2.
GN Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Komatsu T., Suzuki Y., Sugano S.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Suzuki Y., Watanabe M., Sugano S.;
RT "Cloning, expression analysis and chromosome localization of a novel
ACE like enzyme."
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046569; BAB40370.1; -.
DR MEROPS; M02.006; -.
DR InterPro; IPR001548; Peptidase_M2.
DR InterPro; IPR000130; Zn_MTPeptidase.
DR Pfam; PF01401; Peptidase_M2; 1.
DR PRINTS; PR00791; PEPTIDASEA.
DR PRODOM; PD004184; Peptidase_M2; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOMN.1.
SQ SEQUENCE 805 AA; 92491 MW; D2AAB4C2708EB72 CRC64;

Query Match 99.9%; Score 4288; DB 4; Length 805;
Best Local Similarity 99.9%; Pred. No. 5.8e-305;
Matches 804; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 M S S S W L L S L V A V T A A S T I E Q A K T F L D K F N H E A E D L F Y O S S L A S M N Y N T N I T E E N V Q 60
   1 M S S S W L L S L V A V T A A S T I E Q A K T F L D K F N H E A E D L F Y O S S L A S M N Y N T N I T E E N V Q 60
Db 1 M S S S W L L S L V A V T A A S T I E Q A K T F L D K F N H E A E D L F Y O S S L A S M N Y N T N I T E E N V Q 60
QY 61 M N N A G D K W S A F L K E Q S T L A Q M Y P L O E I Q N L T V K L O L A L O Q S S V L S D K S K R L N T I L 120
   61 M N N A G D K W S A F L K E Q S T L A Q M Y P L O E I Q N L T V K L O L A L O Q S S V L S D K S K R L N T I L 120
Db 61 M N N A G D K W S A F L K E Q S T L A Q M Y P L O E I Q N L T V K L O L A L O Q S S V L S D K S K R L N T I L 120
QY 121 N T M S T I Y S T G K V C N P D N P O E C L L E P G L N E I M A N S L D Y N E R L A M E S W S E V G K O L R P L Y 180
   121 N T M S T I Y S T G K V C N P D N P O E C L L E P G L N E I M A N S L D Y N E R L A M E S W S E V G K O L R P L Y 180
Db 121 N T M S T I Y S T G K V C N P D N P O E C L L E P G L N E I M A N S L D Y N E R L A M E S W S E V G K O L R P L Y 180
QY 181 E E Y V L K N E M A R A N H Y E D Y G D Y W R G D Y E V N G V D Y S R G O L I E D V E H T F E E I K P L Y E H L 240
   181 E E Y V L K N E M A R A N H Y E D Y G D Y W R G D Y E V N G V D Y S R G O L I E D V E H T F E E I K P L Y E H L 240
Db 181 E E Y V L K N E M A R A N H Y E D Y G D Y W R G D Y E V N G V D Y S R G O L I E D V E H T F E E I K P L Y E H L 240
QY 241 H A Y V R A K L M N A P S Y I S P I G C L P A H L D M G R F W T N L Y S L T V P G Q K P N I D Y T D A M V Q 300
   241 H A Y V R A K L M N A P S Y I S P I G C L P A H L D M G R F W T N L Y S L T V P G Q K P N I D Y T D A M V Q 300
Db 241 H A Y V R A K L M N A P S Y I S P I G C L P A H L D M G R F W T N L Y S L T V P G Q K P N I D Y T D A M V Q 300
QY 301 A M D A O R I F E A K E F F V S V G L P N M T O G F W E N S M L T D P G N V O K A V C H P T A M D L G K G D F R I L M 360
   301 A M D A O R I F E A K E F F V S V G L P N M T O G F W E N S M L T D P G N V O K A V C H P T A M D L G K G D F R I L M 360
Db 301 A M D A O R I F E A K E F F V S V G L P N M T O G F W E N S M L T D P G N V O K A V C H P T A M D L G K G D F R I L M 360
QY 361 C T K V T M D E F L T A H N E M G H I O Y D M A Y A A O P F L L R N G A N E G F H E A V E I N S L S A A P R K H L S 420
   361 C T K V T M D E F L T A H N E M G H I O Y D M A Y A A O P F L L R N G A N E G F H E A V E I N S L S A A P R K H L S 420
Db 361 C T K V T M D E F L T A H N E M G H I O Y D M A Y A A O P F L L R N G A N E G F H E A V E I N S L S A A P R K H L S 420
QY 421 I G L S P D F O E D N E T E I N F L K Q A L T I V G L P T Y M L E K W M V F E G E I P R D O M K K W M E M 480
   421 I G L S P D F O E D N E T E I N F L K Q A L T I V G L P T Y M L E K W M V F E G E I P R D O M K K W M E M 480
Db 421 I G L S P D F O E D N E T E I N F L K Q A L T I V G L P T Y M L E K W M V F E G E I P R D O M K K W M E M 480
QY 481 K R E I V G V E P V P H D E T Y C D P A S L F H V S N D Y S F I R Y T R T L Y O F O Q A L C O A K H E G P L H 540
   481 K R E I V G V E P V P H D E T Y C D P A S L F H V S N D Y S F I R Y T R T L Y O F O Q A L C O A K H E G P L H 540
Db 481 K R E I V G V E P V P H D E T Y C D P A S L F H V S N D Y S F I R Y T R T L Y O F O Q A L C O A K H E G P L H 540
QY 541 K C D I S N S T E A G Q K L F N M L R L G K S E P W T L A L E N V G A K N M N V R P L I N Y E P L E T W L K D O N K 600
   541 K C D I S N S T E A G Q K L F N M L R L G K S E P W T L A L E N V G A K N M N V R P L I N Y E P L E T W L K D O N K 600
Db 541 K C D I S N S T E A G Q K L F N M L R L G K S E P W T L A L E N V G A K N M N V R P L I N Y E P L E T W L K D O N K 600
QY 601 N S F V G M S T D M S P Y A D O S I K V R I S L K S A L G D K A Y E M N D N E M Y L F R S S V A Y A M R Q Y L K A V N 660
   601 N S F V G M S T D M S P Y A D O S I K V R I S L K S A L G D K A Y E M N D N E M Y L F R S S V A Y A M R Q Y L K A V N 660
Db 601 N S F V G M S T D M S P Y A D O S I K V R I S L K S A L G D K A Y E M N D N E M Y L F R S S V A Y A M R Q Y L K A V N 660
QY 661 Q M I L F G E E D V R A N L K P R I S F N F V T A P K N V S D I I P R T E V E K A I R M S R S R I N D A F R L N D N 720
   661 Q M I L F G E E D V R A N L K P R I S F N F V T A P K N V S D I I P R T E V E K A I R M S R S R I N D A F R L N D N 720

```

Db 661 QMIFGEEDVAVANLKPRISFNFEVTAAPKNVSDIIPREVEKAIKMSRSRINDAFRLND 720
 QY 721 SLEFLGIQPTLGPNNOPPVSIWLVFGVYVIGVILVIFGIDRKKKKRARGSENP 780
 Db 721 SLEFLGIQPTLGPNNOPPVSIWLVFGVYVIGVILVIFGIDRKKKKRARGSENP 780
 QY 781 YASIDISKGENNPGFQNTDDVQTSF 805
 Db 781 YASIDISKGENNPGFQNTDDVQTSF 805

RESULT 3

Q9UFZ6 PRELIMINARY: PRT: 804 AA.
 ID 09UFZ6
 AC 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical 92.3 kDa protein (Fragment).
 GN DKEFZ43A014.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OC NCBL_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=TESTIS;
 RA Mamblatt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL110224; CAB53682.1; -
 DR MEROPS: M02_006; -
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR001130; Zn_MTPeptide.
 DR Pfam: PF01401; Peptidase_M2; 1.
 DR PRINTS: PR00791; PEPTIDPTASEA.
 DR ProDom: PD004184; Peptidase_M2; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 804 AA; 92340 MW; 91FF391074CB5DA9 CRC64;

Query Match 99.8%; Score 4281; DB 4; Length 804;
 Best Local Similarity 99.9%; Pred. No. 1.9e-304;
 Matches 803; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSSSWLLSLVAVTAAGSTIEQATFELDKFNHEAEDLFYQSSIASWNTNTITEENVQ 61
 Db 1 SSSSWLLSLVAVTAAGSTIEQATFELDKFNHEAEDLFYQSSIASWNTNTITEENVQ 60
 QY 62 MNNAAGDKMSAFLEQOSTLAQMYPIQEIQNTLVKLOLALQONGSSVYSEKSKRLNTLN 121
 Db 61 MNNAAGDKMSAFLEQOSTLAQMYPIQEIQNTLVKLOLALQONGSSVYSEKSKRLNTLN 120
 QY 122 TNSTIYSGKVCNPNPQECILLPEGLNEIMANSIDYERLWAMESRSEYKGLRPLYE 181
 Db 121 TNSTIYSGKVCNPNPQECILLPEGLNEIMANSIDYERLWAMESRSEYKGLRPLYE 180
 QY 182 EYVVLKNEMARANHYEDYGDYWRGDEYVNGVDGYDSRGQLLEDVETFEELKPLYEHL 241
 Db 181 EYVVLKNEMARANHYEDYGDYWRGDEYVNGVDGYDSRGQLLEDVETFEELKPLYEHL 240
 QY 242 AYVRAKLMANYPSTISPTGCLPAHLGLDMGREFWNTLVSLVYFEGOKPNIIDVTDAWVQA 301
 Db 241 AYVRAKLMANYPSTISPTGCLPAHLGLDMGREFWNTLVSLVYFEGOKPNIIDVTDAWVQA 300
 QY 302 WDAORIFKEAEKFFVSYGLNNTQGFENSMILDPGNVOKAVCHPAMPDLGKGFRLIMC 361
 Db 301 WDAORIFKEAEKFFVSYGLNNTQGFENSMILDPGNVOKAVCHPAMPDLGKGFRLIMC 360
 QY 362 TKVIMDDFLTAHHEMGHIQYDMAAFAOPFLLRNGANGFHEAVGELIMSLSAATPKLKST 421
 Db 361 TKVIMDDFLTAHHEMGHIQYDMAAFAOPFLLRNGANGFHEAVGELIMSLSAATPKLKST 420

QY 422 GLSPDFQEDNETEINFLKQALTIYGTLPFTYMLEKRMVYKGEIRPKDQMKKRWEMK 481
 Db 421 GLSPDFQEDNETEINFLKQALTIYGTLPFTYMLEKRMVYKGEIRPKDQMKKRWEMK 480
 QY 482 REIVGVAPPHDFTYCDPASPILFVNSDYSEFIRYRTTLYQFOFOBALCOAAKHESPLK 541
 Db 481 REIVGVAPPHDFTYCDPASPILFVNSDYSEFIRYRTTLYQFOFOBALCOAAKHESPLK 540
 QY 542 CDISNSTEAGCKLFNMLRLGKSEPTLALENVGAKMANNRPILNFEPLTKKOONK 601
 Db 541 CDISNSTEAGCKLFNMLRLGKSEPTLALENVGAKMANNRPILNFEPLTKKOONK 600
 QY 602 SFVWGSTDMSPPADOSTIVRISLSKALDQKAYENNNDNMLFRSSVAYAMROYFLKYNQ 660
 Db 601 SFVWGSTDMSPPADOSTIVRISLSKALDQKAYENNNDNMLFRSSVAYAMROYFLKYNQ 650
 QY 662 MIFGEEDVAVANLKPRISFNFEVTAAPKNVSDIIPREVEKAIKMSRSRINDAFRLDND 721
 Db 661 MIFGEEDVAVANLKPRISFNFEVTAAPKNVSDIIPREVEKAIKMSRSRINDAFRLDND 720
 QY 722 LEFLGIQPTLGPNNOPPVSIWLVFGVYVIGVILVIFGIDRKKKKRARGSENP 781
 Db 721 LEFLGIQPTLGPNNOPPVSIWLVFGVYVIGVILVIFGIDRKKKKRARGSENP 780
 QY 782 ASIDISKGENNPGFQNTDDVQTSF 805
 Db 781 ASIDISKGENNPGFQNTDDVQTSF 804

RESULT 4

Q8R010 PRELIMINARY: PRT: 805 AA.
 ID 08R010
 AC 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE RIKEN cDNA 2010305L05 gene.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBL_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC026801; AAH26801.1; -
 KW SEQUENCE 805 AA; 92367 MW; D8B883AAC966A8D9 CRC64;

Query Match 83.4%; Score 3579; DB 11; Length 805;
 Best Local Similarity 82.1%; Pred. No. 4.2e-253;
 Matches 661; Conservative 60; Mismatches 84; Indels 0; Gaps 0;

QY 1 MSSSWLLSLVAVTAAGSTIEQATFELDKFNHEAEDLFYQSSIASWNTNTITEENVQ 60
 Db 1 MSSSWLLSLVAVTAAGSTIEQATFELDKFNHEAEDLFYQSSIASWNTNTITEENVQ 60
 QY 61 MNNAAGDKMSAFLEQOSTLAQMYPIQEIQNTLVKLOLALQONGSSVYSEKSKRLNTLN 120
 Db 61 MNNAAGDKMSAFLEQOSTLAQMYPIQEIQNTLVKLOLALQONGSSVYSEKSKRLNTLN 120
 QY 121 TNSTIYSGKVCNPNPQECILLPEGLNEIMANSIDYERLWAMESRSEYKGLRPLYE 180
 Db 121 TNSTIYSGKVCNPNPQECILLPEGLNEIMANSIDYERLWAMESRSEYKGLRPLYE 180
 QY 181 EYVVLKNEMARANHYEDYGDYWRGDEYVNGVDGYDSRGQLLEDVETFEELKPLYEHL 240
 Db 181 EYVVLKNEMARANHYEDYGDYWRGDEYVNGVDGYDSRGQLLEDVETFEELKPLYEHL 240
 QY 241 AYVRAKLMANYPSTISPTGCLPAHLGLDMGREFWNTLVSLVYFEGOKPNIIDVTDAWVQA 300
 Db 241 AYVRAKLMANYPSTISPTGCLPAHLGLDMGREFWNTLVSLVYFEGOKPNIIDVTDAWVQA 300

QY 301 AMDAQRIFKEAEKFFVSVGLPNNTOGFHNSMLTDPGNVOKAVCHPTAMDICKGDFRILM 360
 Db 301 GMDARIFFOEAEKFFVSVGLPNNTOGFHNSMLTDPGNVOKAVCHPTAMDICKGDFRILM 360
 QY 361 CTKYTMDFLTAHEMGHIQYDMAVAAROPFLRNGANECEFHAVEIGMSLSAATPKHLKS 420
 Db 361 CTKYTMDFLTAHEMGHIQYDMAVAAROPFLRNGANECEFHAVEIGMSLSAATPKHLKS 420
 QY 421 IGLISPDFOEDNTEINFLKQALITVGLTPTMLEKRMVAFGEIPKQMMKKWEM 480
 Db 421 IGLISPDFOEDNTEINFLKQALITVGLTPTMLEKRMVAFGEIPKQMMKKWEM 480
 QY 481 KREIVGVPEVPHDETCYCPASLFHVSNDYSFIRYRTLYOFOQOALCOAAKHEGPLH 540
 Db 481 KREIVGVPEVPHDETCYCPASLFHVSNDYSFIRYRTLYOFOQOALCOAAKHEGPLH 540
 QY 541 KCDISNSTEAGOKLKMSTLGNSEPWTKALENVGARMMDKPLNLYFOPLFLKQNR 600
 Db 541 KCDISNSTEAGOKLKMSTLGNSEPWTKALENVGARMMDKPLNLYFOPLFLKQNR 600
 QY 601 NSFVGMTWSPYADOSTIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVN 660
 Db 601 NSFVGMTWSPYADOSTIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVN 660
 QY 661 QMILFGEEDVRYANLKPRISFNFTAPKNVSDIIPTEVEKAIKMSRSRINDAFRLND 720
 Db 661 QMILFGEEDVRYANLKPRISFNFTAPKNVSDIIPTEVEKAIKMSRSRINDAFRLND 720
 QY 721 SLEFLGIQPTLGPNNOPPVSVIWLVEGVGVIVGIVLITFTGIRDRKKKARSGENP 780
 Db 721 SLEFLGIQPTLGPNNOPPVSVIWLVEGVGVIVGIVLITFTGIRDRKKKARSGENP 780
 QY 781 YASIDIKGENNPGFOFTDDVOTSF 805
 Db 781 YASIDIKGENNPGFOFTDDVOTSF 805
 QY 781 YSDMDICKGESNAGFQNSDADQTSF 805
 Db 781 YSDMDICKGESNAGFQNSDADQTSF 805

RESULT 5

099N71

PRELIMINARY: PRT: 798 AA.

AC 099N71; 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 RT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Angiotensin-converting enzyme-related carboxypeptidase.
 GN 2010305105RIK OR ACE2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Komatsu T., Sugano S., Imai J., Suzuki Y., Hanaoka K., Ymada Y.,
 Hida M., Tanigami A., Muroi S.;
 RT "Molecular cloning, mRNA expression, and chromosomal localization of
 mouse Angiotensin-converting Enzyme-Related Carboxypeptidase.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB053181; BAB40431.1;
 DR MEROPS: M02.006;
 DR MGD: MGI:191728; 2010305105RIK.
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR001680; WD40.
 DR InterPro: IPR00130; Zn_MTPeptide.
 DR Pfam: PF01401; Peptidase_M2; 1.
 DR PRINTS: PR00791; Peptidase_M2; 1.
 DR Prodom: PD004184; Peptidase_M2; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 KW Carboxypeptidase.
 SEQUENCE 798 AA; 91943 MW; 403AEE29D55725A4 CRC64;

Query Match 81.2%; Score 3485; DB 11; Length 798;
 Best Local Similarity 82.2%; Pred. No. 3,1e-246;

Matches 644; Conservative 58; Mismatches 81; Indels 0; Gaps 0;
 QY 1 MSSSSWLLSLVAVTAAGSTIEQAKTFLDKRNHAEDELFOSSLSASNNVTNTEENVO 60
 Db 1 MSSSSWLLSLVAVTAAGSTIEQAKTFLDKRNHAEDELFOSSLSASNNVTNTEENVO 60
 QY 61 MNNAAGDKSAFLKEOSTLAQMYPLQEIQNTLVKQLQALQOQSSSVLSEDKSRILTL 120
 Db 61 MNNAAGDKSAFLKEOSTLAQMYPLQEIQNTLVKQLQALQOQSSSVLSEDKSRILTL 120
 QY 121 NTMSTIYTGKCNPDNFOECLELLEPGNETMANSIDYNEULAMWESRSEVGKOLRPLX 180
 Db 121 NTMSTIYTGKCNPDNFOECLELLEPGNETMANSIDYNEULAMWESRSEVGKOLRPLX 180
 QY 181 EYVVLKNNEMARANNEDYGYWRGDYEVNGVDYDSRQQLDEYVHTFEELKPYEHL 240
 Db 181 EYVVLKNNEMARANNEDYGYWRGDYEVNGVDYDSRQQLDEYVHTFEELKPYEHL 240
 QY 241 HAYVRAKLMAVPSYISPIGCLPAHLIGDMKGRNTMLYSLYFSGQRPNIIDTAMVQ 300
 Db 241 HAYVRAKLMAVPSYISPIGCLPAHLIGDMKGRNTMLYSLYFSGQRPNIIDTAMVQ 300
 QY 301 AMDAQRIFKEAEKFFVSVGLPNNTOGFHNSMLTDPGNVOKAVCHPTAMDICKGDFRILM 360
 Db 301 GMDARIFFOEAEKFFVSVGLPNNTOGFHNSMLTDPGNVOKAVCHPTAMDICKGDFRILM 360
 QY 361 CTKYTMDFLTAHEMGHIQYDMAVAAROPFLRNGANECEFHAVEIGMSLSAATPKHLKS 420
 Db 361 CTKYTMDFLTAHEMGHIQYDMAVAAROPFLRNGANECEFHAVEIGMSLSAATPKHLKS 420
 QY 421 IGLISPDFOEDNTEINFLKQALITVGLTPTMLEKRMVAFGEIPKQMMKKWEM 480
 Db 421 IGLISPDFOEDNTEINFLKQALITVGLTPTMLEKRMVAFGEIPKQMMKKWEM 480
 QY 481 KREIVGVPEVPHDETCYCPASLFHVSNDYSFIRYRTLYOFOQOALCOAAKHEGPLH 540
 Db 481 KREIVGVPEVPHDETCYCPASLFHVSNDYSFIRYRTLYOFOQOALCOAAKHEGPLH 540
 QY 541 KCDISNSTEAGOKLKMSTLGNSEPWTKALENVGARMMDKPLNLYFOPLFLKQNR 600
 Db 541 KCDISNSTEAGOKLKMSTLGNSEPWTKALENVGARMMDKPLNLYFOPLFLKQNR 600
 QY 601 NSFVGMTWSPYADOSTIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVN 660
 Db 601 NSFVGMTWSPYADOSTIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVN 660
 QY 661 QMILFGEEDVRYANLKPRISFNFTAPKNVSDIIPTEVEKAIKMSRSRINDAFRLND 720
 Db 661 QMILFGEEDVRYANLKPRISFNFTAPKNVSDIIPTEVEKAIKMSRSRINDAFRLND 720
 QY 721 SLEFLGIQPTLGPNNOPPVSVIWLVEGVGVIVGIVLITFTGIRDRKKKARSGENP 780
 Db 721 SLEFLGIQPTLGPNNOPPVSVIWLVEGVGVIVGIVLITFTGIRDRKKKARSGENP 780
 QY 781 YAS 783
 Db 781 YDS 783

RESULT 6

099N70

PRELIMINARY: PRT: 353 AA.

AC 099N70; 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 RT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Angiotensin-converting enzyme-related carboxypeptidase.
 GN ACE2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]

SEQUENCE FROM N.A.
 RA Komatsu T., Sugano S., Imai J., Suzuki Y., Hanaoka K., Yamada Y.,
 RA Hida M., Tanigami A., Murai S.,
 RT "Molecular cloning, mRNA expression, and chromosomal localization of
 RT mouse Angiotensin-converting Enzyme-Related Carboxypeptidase";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Komatsu T., Sugano S., Suzuki Y., Hanaoka K., Yamada Y.,
 RT "Molecular cloning of ACE2";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB053182; BAB40432.1; -
 DR MEROPS: M02.006; -
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF01401; Peptidase_M2; 1.
 DR PRINTS: PR00791; PEPTIDPTASEA.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
 DR Carboxypeptidase.
 SQ SEQUENCE 353 AA; 40442 MW; D17B71141EB4AF5B CRC64;

Query Match 35.9%; Score 1539; DB 11; Length 353;
 Best Local Similarity 81.2%; Pred. No. 2.1e-104;
 Matches 285; Conservative 22; Mismatches 44; Indels 0; Gaps 0;

QY 1 MSSSSMLLSLVAVYTAOSTIEQAKTFLDKFNHEDLYQSSLASMNNTNTEENQ 60
 DB 1 MSSSSMLLSLVAVYTAOSTIEQAKTFLDKFNHEDLYQSSLASMNNTNTEENQ 60
 QY 61 MNANNGKWSAFLKQSTLAQMPLOEIONLTVKLOALOONGSSVLSDEKSKRLNTIL 120
 DB 61 MNANNGKWSAFLKQSTLAQMPLOEIONLTVKLOALOONGSSVLSDEKSKRLNTIL 120
 QY 121 NMSTYTGKVCNPDNPOECLELLEPGLEIMANSLDYNERLMAWMSRSEVQKOLRPLY 180
 DB 121 NMSTYTGKVCNPDNPOECLELLEPGLEIMANSLDYNERLMAWMSRSEVQKOLRPLY 180
 QY 121 NMSTYTGKVCNPDNPOECLELLEPGLEIMANSLDYNERLMAWMSRSEVQKOLRPLY 180
 DB 121 NMSTYTGKVCNPDNPOECLELLEPGLEIMANSLDYNERLMAWMSRSEVQKOLRPLY 180
 QY 181 EEEVVLKKNMARNHEDYGDYRGDYEYNGVDGYSRGOLIEDEHTEFEIKPLYEHL 240
 DB 181 EEEVVLKKNMARNHEDYGDYRGDYEYNGVDGYSRGOLIEDEHTEFEIKPLYEHL 240
 QY 181 EEEVVLKKNMARNHEDYGDYRGDYEYNGVDGYSRGOLIEDEHTEFEIKPLYEHL 240
 DB 181 EEEVVLKKNMARNHEDYGDYRGDYEYNGVDGYSRGOLIEDEHTEFEIKPLYEHL 240
 QY 241 HAYVRAKLNAAVPSYSPICGLPAHLIDGMGRFNTNLSLTPPGQKPNIDVTDAVQ 300
 DB 241 HAYVRAKLNAAVPSYSPICGLPAHLIDGMGRFNTNLSLTPPGQKPNIDVTDAVQ 300
 QY 301 AMDAQRIFEAEKFEVSVGLPMTQGFENSMULTDPGNQKAVCHPTAMD 351
 DB 301 AMDAQRIFEAEKFEVSVGLPMTQGFENSMULTDPGNQKAVCHPTAMD 351
 QY 301 GMDAERIFEAEKFEVSVGLPMTQGFENSMULTDPGNQKAVCHPTAMD 351
 DB 301 GMDAERIFEAEKFEVSVGLPMTQGFENSMULTDPGNQKAVCHPTAMD 351

RESULT 7
 Q9GLN6 PRELIMINARY; PRT; 732 AA.
 ID Q9GLN6
 AC Q9GLN6
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE Dipeptidyl carboxy peptidase 1 testicular form.
 GN DCP1.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 NCBI_TaxID=9598;
 RX MEDLINE=20469400; PubMed=11013071;
 RA Dufour C., Casane D., Denton D., Wickings J., Corvol P.,
 RA Jeunemaitre X.;
 RT "Human-Chimpanzee DNA sequence variation in the four major genes of
 RT the renin-angiotensin system.";
 RL Genomics 69:14-26(2000).
 DR EMBL: AF193473; AAC31359.1; JOINED.
 DR EMBL: AF193473; AAC31359.1; JOINED.

DR EMBL: AF193474; AAC31359.1; JOINED.
 DR EMBL: AF193475; AAC31359.1; JOINED.
 DR EMBL: AF193476; AAC31359.1; JOINED.
 DR EMBL: AF193477; AAC31359.1; JOINED.
 DR EMBL: AF193478; AAC31359.1; JOINED.
 DR EMBL: AF193479; AAC31359.1; JOINED.
 DR EMBL: AF193480; AAC31359.1; JOINED.
 DR EMBL: AF193481; AAC31359.1; JOINED.
 DR EMBL: AF193482; AAC31359.1; JOINED.
 DR EMBL: AF193483; AAC31359.1; JOINED.
 DR EMBL: AF193484; AAC31359.1; JOINED.
 DR EMBL: AF193485; AAC31359.1; JOINED.
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF01401; Peptidase_M2; 1.
 DR PRINTS: PR00791; PEPTIDPTASEA.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 732 AA; 83428 MW; C4306443A47E74B CRC64;

Query Match 31.1%; Score 1336; DB 6; Length 732;
 Best Local Similarity 41.6%; Pred. No. 4.3e-89;
 Matches 258; Conservative 119; Mismatches 205; Indels 38; Gaps 10;

QY 15 TAAOS-----TTEQAKTFLDKFNHEDLYQSSLASMNNTNTEB-----NWOM 62
 DB 15 TAAOS-----TTEQAKTFLDKFNHEDLYQSSLASMNNTNTEB-----NWOM 62
 QY 63 NNAQKWSAFLKQSTLAQMPLOEIONLTVKLOALOONGSSVLSDEKSKRLNTIL 122
 DB 63 NNAQKWSAFLKQSTLAQMPLOEIONLTVKLOALOONGSSVLSDEKSKRLNTIL 122
 QY 121 NHT-----LKGTQAKRFVDVNOQNTTKRIKKVQDLEALRALPAQLEENKILTD 172
 DB 121 NHT-----LKGTQAKRFVDVNOQNTTKRIKKVQDLEALRALPAQLEENKILTD 172
 QY 123 MSTYTGKVCNPDNPOECLELLEPGLEIMANSLDYNERLMAWMSRSEVQKOLRPLYE 182
 DB 123 MSTYTGKVCNPDNPOECLELLEPGLEIMANSLDYNERLMAWMSRSEVQKOLRPLYE 182
 QY 173 METTYSVAVCTHNG--SCLQLEPDLTVNMAVTSKRYEDLLMAWGMGKGRALLOFPK 230
 DB 173 METTYSVAVCTHNG--SCLQLEPDLTVNMAVTSKRYEDLLMAWGMGKGRALLOFPK 230
 QY 183 YVVLKKNMARNHEDYGDYRGDYEYNGVDGYSRGOLIEDEHTEFEIKPLYEHL 242
 DB 183 YVVLKKNMARNHEDYGDYRGDYEYNGVDGYSRGOLIEDEHTEFEIKPLYEHL 242
 QY 231 YVELINQARLNGYVDAGSMRSMETPSLE-----QDLERLFOELQPLYLHA 280
 DB 231 YVELINQARLNGYVDAGSMRSMETPSLE-----QDLERLFOELQPLYLHA 280
 QY 243 YVRAKLNAAVPSYSPICGLPAHLIDGMGRFNTNLSLTPPGQKPNIDVTDAVQ 301
 DB 243 YVRAKLNAAVPSYSPICGLPAHLIDGMGRFNTNLSLTPPGQKPNIDVTDAVQ 301
 QY 281 YVRAKLNAAVPSYSPICGLPAHLIDGMGRFNTNLSLTPPGQKPNIDVTDAVQ 340
 DB 281 YVRAKLNAAVPSYSPICGLPAHLIDGMGRFNTNLSLTPPGQKPNIDVTDAVQ 340
 QY 302 WDAQRIFEAEKFEVSVGLPMTQGFENSMULTDPGNQKAVCHPTAMD 360
 DB 302 WDAQRIFEAEKFEVSVGLPMTQGFENSMULTDPGNQKAVCHPTAMD 360
 QY 341 WTPRRMFEKADDFETSLGILPVPPEFWNMSLEKFTDGEVYCHASANDFTNGADFRKQ 400
 DB 341 WTPRRMFEKADDFETSLGILPVPPEFWNMSLEKFTDGEVYCHASANDFTNGADFRKQ 400
 QY 361 CTYVMDDELTAHHEMGIQYDMAYAAOPFLIRNANGEGFHEAVGEISLSAATPKILKS 420
 DB 361 CTYVMDDELTAHHEMGIQYDMAYAAOPFLIRNANGEGFHEAVGEISLSAATPKILKS 420
 QY 401 CTYVMDDELTAHHEMGIQYDMAYAAOPFLIRNANGEGFHEAVGEISLSAATPKILKS 460
 DB 401 CTYVMDDELTAHHEMGIQYDMAYAAOPFLIRNANGEGFHEAVGEISLSAATPKILKS 460
 QY 421 IGLSPDFQDNTEINFLIKQALTVGLPTFLYALIEKRWAVFGELIPKQOMKKNEM 480
 DB 421 IGLSPDFQDNTEINFLIKQALTVGLPTFLYALIEKRWAVFGELIPKQOMKKNEM 480
 QY 461 LNLISSEGSQD--PHDINFLIKMALDKIAFIPIPSYLDQRMWVEFGSTIKENYDQEMSL 519
 DB 461 LNLISSEGSQD--PHDINFLIKMALDKIAFIPIPSYLDQRMWVEFGSTIKENYDQEMSL 519
 QY 481 KRELIVGVEVPEDEYTCOPASLFHVSNDYSPFRYTRILYOFQFQALCAQAHEGPIH 540
 DB 481 KRELIVGVEVPEDEYTCOPASLFHVSNDYSPFRYTRILYOFQFQALCAQAHEGPIH 540
 QY 520 RLKYGGLCPVPPTQGFDPGAKFHIPSSVPIRYIVSVILIOFQFQALCAQAHEGPIH 579
 DB 520 RLKYGGLCPVPPTQGFDPGAKFHIPSSVPIRYIVSVILIOFQFQALCAQAHEGPIH 579
 QY 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVAGAKNNVRPLNPPETLTKDKQK 600
 DB 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVAGAKNNVRPLNPPETLTKDKQK 600
 QY 580 KCDIYQKSGKQGLATAMKLGFSRPMPEAMOLITGPNMSASAMLSYKPLDMLRTENE 639
 DB 580 KCDIYQKSGKQGLATAMKLGFSRPMPEAMOLITGPNMSASAMLSYKPLDMLRTENE 639
 QY 601 --NSFVGW--STDWSPYADQS 617
 DB 601 --NSFVGW--STDWSPYADQS 617
 QY 640 LHGEKLGWPOYMTNPSARS 659
 DB 640 LHGEKLGWPOYMTNPSARS 659

RESULT 8
 Q9GLN7 PRELIMINARY; PRT; 1304 AA.
 ID Q9GLN7
 AC Q9GLN7;
 Q9GLN7;

```

DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE Dipeptidyl carboxy peptidase 1.
GN Dcpl1.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20469400; PubMed=11013071;
RA Dufour C., Casane D., Denton D., Wickings J., Corvol P.,
RT Jeunemaitre X.;
RT "Human-Chimpanzee DNA sequence variation in the four major genes of
RT the renin angiotensin system.";
RL Genomics 69:14-26 (2000)
DR EMBL; AF193462; AAC31358.1; -
DR EMBL; AF193463; AAC31358.1; JOINED.
DR EMBL; AF193464; AAC31358.1; JOINED.
DR EMBL; AF193465; AAC31358.1; JOINED.
DR EMBL; AF193466; AAC31358.1; JOINED.
DR EMBL; AF193467; AAC31358.1; JOINED.
DR EMBL; AF193468; AAC31358.1; JOINED.
DR EMBL; AF193469; AAC31358.1; JOINED.
DR EMBL; AF193470; AAC31358.1; JOINED.
DR EMBL; AF193471; AAC31358.1; JOINED.
DR EMBL; AF193472; AAC31358.1; JOINED.
DR EMBL; AF193473; AAC31358.1; JOINED.
DR EMBL; AF193474; AAC31358.1; JOINED.
DR EMBL; AF193475; AAC31358.1; JOINED.
DR EMBL; AF193476; AAC31358.1; JOINED.
DR EMBL; AF193477; AAC31358.1; JOINED.
DR EMBL; AF193478; AAC31358.1; JOINED.
DR EMBL; AF193479; AAC31358.1; JOINED.
DR EMBL; AF193480; AAC31358.1; JOINED.
DR EMBL; AF193481; AAC31358.1; JOINED.
DR EMBL; AF193482; AAC31358.1; JOINED.
DR EMBL; AF193483; AAC31358.1; JOINED.
DR EMBL; AF193484; AAC31358.1; JOINED.
DR EMBL; AF193485; AAC31358.1; JOINED.
DR MEROPS; M02.001; -.
DR InterPro; IPR001548; Peptidase_M2.
DR InterPro; IPR000130; Zn_MTEpeptidse.
DR Pfam; PF01401; Peptidase_M2; 2.
DR PRINTS; PR00791; PEPDIPASEA.
DR ProDom; PD004184; Peptidase_M2; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN.1.
SQ SEQUENCE 1304 AA; 149369 MW; DCF728D0BA0F1314 CRC64;

Query Match
Best Local Similarity 31.0%; Score 1329; DB 6; Length 1304;
Matches 254; Conservative 118; Mismatches 205; Indels 34; Gaps 9;

DY 20 TIEQAKTFLDKFNHEADLFYQSSLASWNTNTTET-----VNOANNAGDKWSA 71
DY 642 TTEAEASKFEVEYDRTSOVWNEVEAMNNTNTTETSKILLQKNQJANHT----- 695
DY 72 FLKEGSTLAOMYPLDIOINLVKLOLQALQONGSSVLESDSKRLNTLNTMTSTYSGK 131
DY 696 --LKGIGQARRFDVNOJLQNTTIKRIKKVQDLERAAALPAQLEBYKNKILLDMETTSVAT 753
DY 132 VCNPNPQECILLPEGLMEIMANSIDYNERLWAMESRSEVGKQRLPYEYVVKNEMA 191
DY 754 VCHTNG--SCQLPEPDLINWATSRKEDLLWAGWMDKAGRIILOFYRYVELINQAA 811
DY 192 RANHEEDGVYKRGDYEVNGVDYDSRGQIIEDEVEFEIKPLYEHLHAAYRAKLMA 251
DY 812 RLNGVNDAGDSWRSRYETPSLE-----QDLERLQELQDLYVNLHAAYRRALHNR 861
DY 252 Y-PSYISPIGCLPAHLGLGDMWGREFWNLVSLTYVFGKRPNIIDVTADAVQADQRIEKE 310

```

```

DB 862 YGQNHINLEGPIPAHLGLGDMWGREFWNLVSLTYVFGKRPNIIDVTADAVQADQRIEKE 921
DY 311 AEKFPYVGLPLMWTGQFWNSMLTDPGNVQKAVCPHTAMDGKG-DERIIMCTKVWDDF 369
DY 922 ADDEFSTSLGLFVPPPEFNKRSMLTEKPTDGRVYVCHASAMDFYCNKDPRIQCTVNLDEL 981
DY 370 LTAHHEMGIQYDMAVYAAQPFLLRNGAGNEHVEGELMSIATPYNHLSTGLSPDFQ 429
DY 982 VVAHHEMGIQYFMOYKPLPALNGAGNPGFHEAIGVLAISVTPHLSLNLSSSEG 1041
DY 430 EDNETEINFLKQALTTVGLTPTFYMLEKWMVAFKEIPRODMKRWMEKREIYGVY 489
DY 1042 SD-EHDINFLMKALDKIAFIPFSYLDQWWRVFDSDIIRKENVNDEMWSLRKLYGGLCP 1100
DY 490 PYPHEITYCDPASLPLFVNSDSFRITRYTRLYPQFQDALQAKHGLKCDISNSTE 549
DY 1101 PVPRTQGDPPGAKFHIPSSVPIRYFVSFIQGFHEALCQAAHGTGPLHKCDIYQSK 1160
DY 550 AGQKLFNMLRLKSEPTLALENVYGAKNMNVRLNTPPELFTWLKDONK--NSFVGN- 606
DY 1161 AGQRLATAMKLGFSRPPPEAMQILITGPNNASAMLSYFPLDMLRTENELHGEKLGWP 1220
DY 607 STDMSPPADQS 617
DY 1221 QYNMTPNSARS 1231

RESULT 9
ID 09EQW9 PRELIMINARY; PRT; 1313 AA.
AC 09EQW9;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, last annotation update)
DE Angiotensin-converting enzyme.
GN ACE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F344/N; TISSUE=LUNG;
RA Jafarian-Thrani M., Listwak S., Barrientos R.M., Michaud A.,
RA Corvol P., Sternberg E.M.;
RT "Characterization of a missense mutation in the angiotensin I-
RT converting enzyme cDNA in exudative inflammation resistant F344/N
RT rats.";
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DL EMBL; AF201331; AAC35596.1; -
DL InterPro; IPR001548; Peptidase_M2.
DL InterPro; IPR000130; Zn_MTEpeptidse.
DL Pfam; PF01401; Peptidase_M2; 2.
DL PRINTS; PR00791; PEPDIPASEA.
DL ProDom; PD004184; Peptidase_M2; 2.
DL PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN.1.
SQ SEQUENCE 1313 AA; 150941 MW; E888341562542B04 CRC64;

Query Match
Best Local Similarity 42.0%; Score 1310; DB 11; Length 1313;
Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps 7;

DY 20 TIEQAKTFLDKFNHEADLFYQSSLASWNTNTTET-----VNOANNAGDKWSA 79
DY 650 TDEAKANFEVEYDRTAKVLMNEVEAMNNTNTTETSKILLQKNKESNHTLKYGTW 709
DY 80 AQMPPLDIOINLVKLOLQALQONGSSVLESDSKRLNTLNTMTSTYSGKVCNPDNQ 139
DY 710 AKTFDVSNFQNSTIKRIKKVQVNDRAVLPNELEBYNQILLDMETTSVANVCYTG-- 767
DY 140 ECLLEGLNLEIMANSIDYNERLWAMESRSEVGKQRLPYEYVVKNEMARANHYEDY 199

```

```

Db 768 TCTSLSEPDITINMATSRYEELLWMSKRDVGRALLPEFPKYVDPSNKIANKNGYSOA 827
Oy 200 GDIYRGDEYNGVDGYYSRQGLIEDEYHEFEETKPLYEHLHAYVRAKLMNAYPS-YISP 258
Db 828 GDSRSSSESDLE-----QDLKTYQELQPLYLHAYVRSRSLHRYGSEYINL 877
Oy 259 IGCPLPAHLGDMGREFWNTLSLTPFGOKPNIDVDMDAMQAMDARTEKEKEFPVS 318
Db 878 DGPFAHLGMMQATMSNIYDVAFPSPASIDATEAMIKOGWTPRIKPEADNFTSTL 937
Oy 319 GLPNTQGFWENSMITDGNVQKAVCHPTAMDLSGK-DEFILMCTKYMTDMDFLTAHENG 377
Db 938 GLPVPPEFWMKSMLEKPTDREVVACHASAMDFYNGDFRIKOCSTYVMEELVIAHENG 997
Oy 378 HIOYMAVAAOPELLRNANGFHEAVGEIMSLSATPKHLKSGILSPDQEDNETEIN 437
Db 998 HIOYMQKDLPTVTFREANGFHEALIDVALSVTPKHLHSLNLSSSE-GSGYEHDIN 1056
Oy 438 FLKQALITVGLTLPFYWMLEKRWKVEKGETPKDQMKKWKEMKREIVGVPEPHDET 497
Db 1057 FLKMAALDKIAFIPFSYLDQMRWRFVFDGSLTKENYQEMWSLRLKYQGLCPVPSOGD 1116
Oy 498 CPASLPHNSDYSFIRYRTYTRLYOFQFOALCOAKHEGPLKCDISNSTEAGOKLFNM 557
Db 1117 FPGSKFHPANVPYIRIFISITIOFQHEALCRAAGHTGPLYKCDIYOSKEAGKLLADA 1176
Oy 558 LRLGSEPTWLTALENVGAKNNVRLNTEPELTWLDONK--NSFVGM-STDWSP 612
Db 1177 MKLGYSKQPEAKMKTITGQPNMSASAIMYFKPLETMTLVTEMRHGGELTGWPEYTW 1234

```

RESULT 10

```

Q15540 PRELIMINARY: PRT: 694 AA.
ID 015540:
AC 015540:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DCP1 protein.
GN DCP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90046671; PubMed=2554286;
RA Ehlers M.R., Fox E.A., Strydom D.J., Riordan J.F.;
RT "Molecular cloning of human testicular angiotensin-converting enzyme:
RT the testis isozyme is identical to the C-terminal half of endothelial
RT angiotensin-converting enzyme.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:7741-7745(1989).
DR EMBL; M2658; AAA60612.1;
DR InterPro: IPR001548; Peptidase_M2.
DR InterPro: IPR000130; Zn_MTpeptidase.
DR Pfam: PF01401; Peptidase_M2; 1.
DR PRINTS; PR00791; PEPDPTASEA.
DR PRODOM; PD004184; Peptidase_M2; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN.1.
SQ SEQUENCE 694 AA; 79333 MW; 57C0F9C5A5CA119 CRC64;

```

Query Match 30.4%; Score 1303; DB 4; Length 694;
 Best Local Similarity 42.7%; Pred. No. 1e-86;
 Matches 250; Conservative 107; Mismatches 195; Indels 34; Gaps 9;

```

Oy 45 LASWNTNTITEE-----NVONMNNAGDKMSAFLEKESSTLAQMPLOEITONLVKQ 96
Db 57 IATKYNNTITETSKILQKNOQIANHT-----LKQYQARKFDVNOQNTTKRI 108
Oy 97 LQALQONSSVLSSEKSKRLNTIIMSTIYSGVCNPNPQCLLEGLEINMANSL 156
Db 109 IKKVDLERRALPAQGLEENKILLDMETYSVATVCHPNG--SCDLEPDITINMATS 166

```

```

Oy 157 DYNRLAMWESMSSEVQKOLRPLYEEVYLKNEMARANHEDYGDYRGDEYNGVDGD 216
Db 167 KYEDLWAMEGMRDKARAILQFPKVELLINOAAHLNGYVADAGDSMRMETPSLE--- 223
Oy 217 YSRQGLIEDEYHEFEETKPLYEHLHAYVRAKLMNAY-PSYISPGCLPAHLGDMGREF 275
Db 224 -----QDLERLQELQPLYLHAYVRSRSLHRYGSEYINL 276
Oy 276 TNLVSLVTPFGOKPNIDVDMDAMQAMDARITEKEKEFPVSGLPNTQGFWENSMITD 335
Db 277 SNIVDLVVPFSPASMDTEAMKOGWTPRIKPEADNFTSTLGLLPVPEFWMKSMLEK 336
Oy 336 PGNVQKAVCHPTAMDLSGK-DEFILMCTKYMTDMDFLTAHENGHIQDMAVAAOPELLRN 394
Db 337 PLDGRVYVCHASAMDFYNGDFRIKOCSTYVMEELVIAHENGHIQYMQKDLPAALRE 396
Oy 395 GANEGFHEAVGEIMSLSATPKHLKSGILSPDQEDNETEINFLKQALITVGLTLPFY 454
Db 397 GANPFHEALIGDVIALSVTPKHLHSLNLSSSEGSQ- EHDINFLKMAALDKIAFIPFS 455
Oy 455 MEKRWKVEKGETPKDQMKKWKEMKREIVGVPEPHDETCDPASLPHNSDYSFIR 514
Db 456 LVDQMRWRFVFDGSLTKENYQEMWSLRLKYQGLCPVPSOGDPPGAKPHIPSSVYIR 515
Oy 515 YTRTLYOFQFOALCOAKHEGPLKCDISNSTEAGOKLFNMLRLGSEPTWLTALENV 574
Db 516 YFVSFLIOFQHEALCOAGHTGPLKCDIYOSKEAGRLATAKLGFSPWPEAMQLIT 575
Oy 575 GAKNNVRLNTEPELTWLDONK--NSFVGM-STDWSPYDAS 617
Db 576 GQPNMSASAIMYFKPLETMTLVTEMRHGGELTGWPEYTWNTNSARS 621

```

RESULT 11

```

Q9NDS8 PRELIMINARY: PRT: 648 AA.
ID 09NDS8:
AC 09NDS8:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ecdysteroid-inducible angiotensin-converting enzyme-related gene
DE product.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C108; TISSUE=WING IMAGINAL DISK;
RX MEDLINE=20556050; PubMed=11102839;
RA Quan G.X., Mita K., Okano K., Shimada T., Uga Jin N., Xia Z., Goto N.,
RA Kanke E., Kawasaki H.;
RT "Isolation and expression of the ecdysteroid-inducible angiotensin-
RT converting enzyme-related gene in wing discs of Bombyx mori.";
RL Insect Biochem. Mol. Biol. 31:97-103(2001).
DR EMBL; AB026110; BAA97657.1;
DR MEROPS; M02.002;
DR InterPro: IPR001548; Peptidase_M2.
DR InterPro: IPR000130; Zn_MTpeptidase.
DR Pfam: PF01401; Peptidase_M2; 1.
DR PRINTS; PR00791; PEPDPTASEA.
DR PRODOM; PD004184; Peptidase_M2; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN.1.
SQ SEQUENCE 648 AA; 74917 MW; 9A740AA9FCACEF0 CRC64;

```

Query Match 25.4%; Score 1089.5; DB 5; Length 648;
 Best Local Similarity 35.6%; Pred. No. 3.9e-71;
 Matches 232; Conservative 125; Mismatches 259; Indels 35; Gaps 11;

```

Oy 11 LVAVTAOSTIEQAK-TFLDKENHEADLFT-----QSLASWNTNTITEEN 58
Db 10 LIAAIVAVFIVATQGRDPLEAREHEAREYMLHLKATYGLRKNRASIADAEWETYSNITKEN 69

```


FT	CHAIN	30	660	POTENTIAL.
50	SEQUENCE	660 AA;	75257 MW;	6F164CF70C938E63 CRC64;
	Query Match	24.6%;	Score 10577;	DB 5; Length 660;
	Best Local Similarity	37.2%;	Pred. No. 9,46-69;	
	Matches 226;	Conservative 110;	Mismatches 237;	Indels 34; Gaps 13.
QY	11	LVAVTAAGSTIEBOAK--TELDKPFHEADELDYQSGSLASNNYNINTEENYQNNKN--AGD	67	
DB	36	LATLSNVSAIKIKDEAGVAFIEGLNDPPTYYTINNVDSSSSWDYASNTIDYN--QNNSKNVST	94	
QY	68	KWSAFIKESQSTLQAMYPLOEIONLTGVKLOLQALQONGSSVSEDEKSKRLNTILNTMSTYI	127	
DB	95	EYVSKMEQGEITAKRRDWMNFKNDSIKRLFRIVATIGLAALRDDKLEMASTLSKMAIY	154	
QY	128	STGKVG--NPDPNPOCLLIEPGLMEIANSIDYVERLAMEWSRSEYKOLRPLEEYV	184	
DB	155	GSTKVFVGKDKLP-----LEPDLFRNKEVENVYDKLLQTLWLANHNAVPAIKOYIYPI	209	
QY	185	VLRKEMARAHYEDYGDYGRGDEYGVNGVGYYSKGLLEDEYHTEFRIKRYLHNAAV	244	
DB	210	KLSEASALSDGYNIKISAMLSDEYE-----NMTELYDKLWEDLSPLVKKLHAIV	259	
QY	245	RAKIMNAYPSYIPIGCLPAHLILGDMKGRFTNLY--SLTVPGQKPNIDYDAVDAQMD	303	
DB	260	RAKIREIYPERLDEDTITAHLLGNMAQDEGTILPHLTME--DKP--LDISKTVQDKWD	316	
QY	304	AQRIFKEAEKFEYVSGILPNNTOGFWMENSLDPGNVOKAVCHPTAAMDIGK--DERIIMCT	362	
DB	317	AQKFNHAEDEFTSLGIDNMTSEFMSKILTPRED--RELQHASAMNNGDDPRIKCT	375	
QY	363	KVTMDFLVNHHEMGIQYDMAAOPRLIRGANGEGNEAVGELMISAATPRHLKSTIG	422	
DB	376	DSPEVLERTYHNHEMGIIEYYMOKYKHLNVLLQGANEGHEAAGDILALSVATKHYEKL	435	
QY	423	LISDPDEDETEINFLKQALITVGTLPFTYLEKRYMVEKGEIRPDQMKFWMEKRR	482	
DB	436	LLEP--TDKYNADVLILASALDKIAFLFFGVLIDKMKWTITPTGTPRDKANEKFWEXRI	492	
QY	483	EIVGVPEVPHDEYCDPASLFEHVSNDYSFIRYRTTYLOFOFQALCOAQ--HEGRL	539	
DB	493	KYQGVSPFVARNSEFPFGGAKYVHALHYLYKRYFAVFILOQFHNILSTVAKKVDENHRE	552	
QY	540	HKCDISNSTERGOKLEMLERLKSSEPWTLALENVUGAKMAYKRLNLFEBELFTWIKDON	599	
DB	553	HECDIYEKNAAGVLLKKGSLGRSKPWPVLEIMAGTRQMSASLSLKKYPERLEKWLDERI	612	
QY	600	KNSFVGW	606	
DB	613	KNEYVGV	619	
RESULT 14				
Q24222	Q24222	PRELIMINARY;	PRT;	630 AA.
AC	Q24222;			
DT	01-NOV-1996 (Tremblrel. 01, Created)			
DT	01-NOV-1996 (Tremblrel. 01, Last sequence update)			
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)			
DE	METALLOPEPTIDASE (Angiotensin-converting enzyme-related protein).			
GN	ACER OR CG10593.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_Taxid=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91128790; PubMed=8973330;			
RA	Taylor C.A.M., Coates D., Shitras A.D.;			
RT	"The Acer gene of Drosophila codes for an angiotensin-converting			
RL	enzyme homologue."			
	Gene 181:191-197(1996).			

```

[2]
SEQUENCE FROM N.A.
RP STRAIN-BERKELEY;
RA Ceiniker S.E., George R.A., Galle R., Swirskas R.R., Hoskins R.A.,
RA Agbayani A., Arcaina T.T., Baxter E., Blazey R.G., Chavez C., Chew M.,
RA Doyle C.M., Farfan D.E., Flanagan J., Houston K.A., Hummasti S.R.,
RA Karia K., Kearney L., Kim S.H., Lee B., Lotman M.A., Mak J.,
RA Mada P., Mok M.S., Moshrefi A.R., Moshrefi M., Nixon R., Pacleb J.M.,
RA Park S., Pfeiffer B., Punch E., Snir E., Twomey B., Wan K.H.,
RA Whitelaw K.R., Yee A., Zhang R., Zieran L.L., Kimmel B.E.,
RL Submitted (MAY-1998) to the EMBL/GenBank/DBD databases.
DR EMBL: X96913; CAA5632.1; -
DR EMBL: AC004728; -; NOT_ANNOTATED_CDS.
DR FlyBase: FBgn001612; Acer.
DR InterPro: IPR001548; Peptidase_M2.
DR InterPro: IPR00130; zn_MTPeptide.
DR Pfam: PF01401; Peptidase_M2; 1.
DR PRINTS: PR00791; PEPTIDPTSEA.
DR PRODOM: PD004184; Peptidase_M2; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
FT CONFLICT 236 L -> F (IN REF. 2).
FT CONFLICT 341 Q -> H (IN REF. 2).
FT CONFLICT 528 V -> A (IN REF. 2).
SQ SEQUENCE 630 AA; 73042 MM; 248AAB90319BC32D CRC64;

Query Match          24.08; Score 1030; DB 5; Length 630;
Best Local Similarity 35.68; Pred. No. 8.3e-67;
Matches 219; Conservative 113; Mismatches 260; Indels 24; Gaps 10;

QY 6 WLLSLVAVTAASOTIEEAKTFLDKFHEADLFYQSSLASNYNTMTTEENVQNMNNA 65
DB 16 WLPGLISMGNCSASYLF-ARRFELENEDELRRRPFHEPLISGNTNTNTNTEARQMIIEV 74
QY 66 GDKWSAFLKQSLTQAOPILOETQNTLVKLOALQOQSSVLSDEKSKRLNTLITMTST 125
DB 75 YARNALNKRLAQOIKSSDYVQSEDADIRQAEHLISKLGSALNADYIALQNAISSMGT 134
QY 126 IYSTGWKCNPDNPEC-LILEPLNLINMNSLDYNRILNAMESWSEVQKOLRPLEEYV 184
DB 135 NYATAVCTSTNNSDCLTEPHIOERLSRDPALAWYWRMWHKSGTPMKONFAEYV 194
QY 185 VLKEMARNHEDYDYGVDYGVNGVGYDYSRGQLLEDYEHPEETKLYEHLHAYV 244
DB 195 RLRRKASQNLGHSYADYVQFE-----DPRFER-----QIDATFTKQLPLRQLHGYV 244
QY 245 RALILNANAY-PSYISPTIGCLPAHLIGDMGRFMTNLYSLTPRGQKNIDVTAMVDQAWD 303
DB 245 RFLRQHYGPDVWAPAGNIPISILGNMGQSWNELDLFTPEPEKRFVYVAKMEKQYGT 304
QY 304 AORIFPEAKRFYSVGLPMNTOGFVNSMLTPRGVQKAVCHPTIADLQK-GDPRILMCT 362
DB 305 VQLEFELGDFQSLGRALPSPFNNLSVLTTRPDD-ROVVCASASAMDFQDSIDVRIKMT 363
QY 363 KYTMDDELTAHHEMGHIQYDMAUYAOPFLLRNGANGFEEAGELINSTISAAPRKLKSG 422
DB 364 EVDSHFYVYVHHELGHIQYLYOYEDQPAVYRARPNGHEAVGDIYALISVSAKTLKAG 423
QY 423 LLSPDFQEDNEELNFKLQALITVGLTPTMYLKEKRWVFKGEIPKQDMKKWMEKMR 482
DB 424 LIE-NGLRDEKSRINQLEFQALSKIVLPEFGYAVDYKRYAVRNLDSQNMCGWQMRBS 482
QY 483 EIVGVVEPRPHDETCODPASLFHVSNDSFITYYRITLYKQFOFQALCOAKHESG----- 538
DB 483 EFGGVPEPRFETEKDFDPKAKYHIDADVEYLYEFAAHIFQFOFHVLRKKAQYAPNNSR 542
QY 539 --LHHCDDISNTEAGOKLFNNMLRLGKSPWTLLENVYGAKNMNRPPLNFEPLFTLK 596
DB 543 LFLDNDODIFGSKAAGRSLSQPLSKNSNWKKEVLEEFGETEMDPAALLEVEPLYLWLK 602
QY 597 DONKNSEVGMSTDMSP 612
DB 603 QE--NSRLGVPLGMGP 616

```

```

RESULT 15
ID Q9VLU6 PRELIMINARY; PRT; 630 AA.
AC Q9VLU6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ACER protein (U08328P).
DE ACER OR G10593.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731133;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Bax A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Dubin K.J., Evangelista C.C., Ferraz C., Gelbart W.M., Glasser K.,
RA Fostier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kension J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattie B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinett K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Spaden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gbbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Ceiniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBD databases.
DR EMBL: AE003621; AAF52693.1; -
DR EMBL: AY051750; AAK93174.1; -
DR MEROPS: M02.002; -
DR FlyBase: FBgn001612; Acer.
DR InterPro: IPR001548; Peptidase_M2.
DR InterPro: IPR00130; zn_MTPeptide.
DR Pfam: PF01401; Peptidase_M2; 1.

```

DR PRINTS: PR00791; PEPTIDPTASEA.
 DR Prodom: PD004184; Peptidase_M2; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN 1.
 SQ SEQUENCE 630 AA; 73057 MW; 6D9355EB5773289 CRC64;

Query Match 24.0%; Score 1028; DB 5; Length 630;
 Best Local Similarity 35.6%; Pred. No. 1.2e-66;
 Matches 219; Conservative 112; Mismatches 261; Indels 24; Gaps 10;

QY 6 WLILSLVAVTAAGSTIEQATFLDKFNHNEDELFTYSSLASAWYNTNTEENVQNNNA 65
 DB 16 WPLHGLSNGSCSASVLE-ARFPELNEQDRLRRFHEEPISGVYNNVTNTEANQAMIEV 74
 QY 66 GDKNSAFLEQSTLAOMPTLOEIONLTVKLOLALQONQSVLSSEDKSKRLNTLTMST 125
 DB 75 YARNAELNKRRLAQOIKSSDYQSDADIRQAEHLKSLGASALNADYIALONAISSMOT 134
 QY 126 IYSTGKVCNPDNPQEC-LLEPGLNEMANSLDYNERLWAMESRSEVQKOLRPLYEEV 184
 DB 135 NYATATVCSTYNRSDCSLTLEPHIOERLSRDPALAWYWRHHDKSGTPMRQNFAYV 194
 QY 185 VLKEMARAHNYEDYGYWRDYEYVNGVDYDSRGQIEDVHTFEETPLTEHHAAY 244
 DB 195 RLTKKASQLNHRSYADYWOYE----DDEFR-----QLDATERQLPFYFOLHGYV 244
 QY 245 RAKIMNAY-PSYISPIGCLPAHLLDGMGRFMTNLSLTVPEGOKPNIDTDAVQDAMD 303
 DB 245 RFRRLRQHYGPDYMAEGNIPISLIGNMWGSNELLDFPYPKRPVDAKAEKOCGYT 304
 QY 304 AQRIFKAKEFEVSVGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDIGK-GDFRILMCT 362
 DB 305 VOKLFELGDOFQSLGRALPPSEFMNLSVLRPDD-RHYVCASAMFYDSDSVRIMCT 363
 QY 363 KVTMDDELTAHHEHGHIOYMAAOPFLRNGANEGFHRVAGEIMLSATKHLKSTIG 422
 DB 364 EVDSHYTYVAHHEGHIOYLYQEQPARYRGADNPGEHVAVGIVIALSVMSAKHLKAIG 423
 QY 423 LLSPEFQEDNETELNFKLQALITVGLPFTYMLEKRWMMVFGEIKEDOMKMMWEMKR 482
 DB 424 LIE-NGRLDEKSRINQFLKQALSKTYFLPGYAVDKRYAVFRRELDSEOMNCFWOMRS 482
 QY 483 EIVGVVEVPVPHDETYCDPASLFEHVSNDYSFIRYTRTLQYQFOEALQOAKHEGP--- 538
 DB 483 EFGGVPEVPVFEETKEDFPKAYHIDADVEYLRFAHIFQEFKALCRKAGCVAPNNSR 542
 QY 539 --LHKCDINSTGCKLFNMLRKGSEPTLAEENVGAKNNVRLVYFPEPLTWK 536
 DB 543 LTLNCDIFGSKAAGRSLSQFLSKGNSRHKVEYLEEFTGETEMDPALLLEYFPEPLYOMLK 602
 QY 597 DQNKSEFGWSTWSP 612
 DB 603 QE--NSRLGVPLGWP 616

RESULT 16
 Q9D836 PRELIMINARY; PRT; 249 AA.
 ID Q9D836
 AC Q9D836;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE 2010305L05Rik protein.
 GN 2010305L05Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE;
 RX MEDLINE=21085660; Pubmed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuo H. A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuenl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L. M., Stabli R., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,
 RA Brownstein M. T., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D. A., Kamuya M., Lee N. H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K. P.,
 RA Suzuki H., Toyokawa K., Wang K. H., Weltz C., Whitaker C., Wilmink L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlschki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 DR EMBL: AK008530; BAB25723.1; -;
 DR MEROPS: M02.006; -;
 DR MGD: MGI:1917258; 2010305L05Rik.
 DR InterPro: IP001548; Peptidase_M2.
 DR Pfam: PF01401; Peptidase_M2; 1.
 SQ SEQUENCE 249 AA; 28379 MW; 19372B2B78AAE921 CRC64;

Query Match 23.4%; Score 1004; DB 11; Length 249;
 Best Local Similarity 75.1%; Pred. No. 1.7e-65;
 Matches 187; Conservative 31; Mismatches 31; Indels 0; Gaps 0;

QY 557 MRLRGSEPTLAEENVGAKNNVRLVYFPEPLTWKLDQNKNSFGWSTWSPYQD 616
 DB 1 MSLGSEPTLAEENVGAKNNVRLVYFPEPLTWKLDQNKNSFGWSTWSPYQD 60
 QY 617 SIKVRLSLKSLGDKAYENNDNMYLFRSSVAYAMRQYFLAKYKNQMLFGEEDVAVANIK 676
 DB 61 SIKVRLSLKSLGDKAYENNDNMYLFRSSVAYAMRQYFLAKYKNQMLFGEEDVAVANIK 676
 QY 677 PRISFEFTAPKNVDIIPREVEKAIMRSRINDARLNDNSTLEPLGIPTGPPNQ 736
 DB 121 PRVSFEFTAPKNVDIIPREVEKAIMRSRINDARLNDNSTLEPLGIPTGPPNQ 180
 QY 737 PVSIMLIFGVGVVIVGIVILFTGIRDRKKRKAKSGENPYASIDISGENNPGFO 796
 DB 181 PPTIMLIFGVVYVGVVIVGIVILFTGIRDRKKRKAKSGENPYASIDISGENNPGFO 796
 QY 797 NTDDVOFTSF 805
 DB 241 NSDDAQTSF 249

RESULT 17
 Q8SXX2 PRELIMINARY; PRT; 844 AA.
 ID Q8SXX2
 AC Q8SXX2;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE RH06639p.
 GN ANCE-3.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Mirande R., Mungall C. J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phohanavong S., Wan K., Yu C., Lewis S. E., Rubin G. M.,
 Celniker S.;

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY075532; AAL68339.1; -
 SO SEQUENCE 844 AA; 98243 MW; 46E1E123F5CC8F71 CRC64;

Query Match 21.7%; Score 930; DB 5; Length 844;
 Best Local Similarity 33.2%; Pred. No. 2,6e-59;
 Matches 194; Conservative 119; Mismatches 248; Indels 24; Gaps 9;

QY 46 ASNYNTNTEENVQNMNNAAGKWSAFLKEQGSNTLQMTYLOEQNTYVLOLQALQNGS 105
 DB 234 AONNFEYVNDFTQTALNAOQRYVEFORITAEQSKRIKDLFDRLRLTQMLQSEVGP 293
 QY 106 SVLSEKSKRLNTILNTMSTIYTGKVCNPDNPOEC-LILEPGLNEIMANSIDYNERLWA 164
 DB 294 NALPLDVBRYNRLNLEMLFVNSAETICAYQOPQCDLHYITQLKIDYAKSDMDELQMTW 353
 QY 165 WESRSEVGKOLRPLYEEVYVLKNEBARNHEDYDYGKDEYVNGVDGYTSRQGLTE 224
 DB 354 WVEYHKKAGGMDSEYQOLIDVQAEVAYVNNVTNGEYWAYE-----SGNFRQ 403
 QY 225 DVEHTFEETKPLYEHLHAYVRKIMAY--PSYISPIGCPAHLIDGMGRFNTNLSLV 283
 DB 404 DMIDYEQILPYLEGLHAYVRKLDYGPDRINRIAPISHILGMVGSNSVLDLIT 463
 QY 284 PEGQKNIDVTDAVDAQDAQRIEKAKEKFEVSVGLPMMTOGFWMENSLTPGVQKAV 343
 DB 464 PYGRKLIDVTPEVNGVGTPLQMLADEEFTSINMSAVGEFRNSLFEQPLD-RVYL 522
 QY 344 CHTPANDL-GKGFRLMCTKYTMDFLTANHEMGIQDYAMAYAOPELLRNGANGEGHE 402
 DB 523 CEPSADDFCNHDFRVKICTDINQSRSLISVHEMAHIQYFLQYRHLPKIFRNGANPAFRQ 582
 QY 403 AVEIMSLSAATPKHLKLSLSPDQENENEFNLQALTYGTFPTMLEKMRM 462
 DB 583 AVDAIGLSVSTPRHLQTLGLQSLDE--SSDINDLFTMALDKYALFALSLDWRD 641
 QY 463 VFKEGILPKDQMKMKWEKREIYGVVEPYPHDEYCDPASLPHVSNDSYFIYRTTLYQ 522
 DB 642 VFGSNANKRPMCHNYNLEKEKSGIKRPLRSEKDFDPGAKHIANIPIYIFFSTVLD 701
 QY 523 FQEQEALCOAAK-----EGPLHKCDISNSTEAGOKLENNMLRLKSEPTWTLALENVGA 576
 DB 702 FOYLRGLCRBSGQYVGPDRPKPLHOCDIYROPAGNLTKLMSKASQPMQOELEETLRE 761
 QY 577 KNNVAPRLNLYFEPLFTMKDON--KNSFVGSTDMSPADYSIX 619
 DB 762 GRLDGTALREYFALEEMLRQEMLRITNEYGMNYP-GDYCKRSIE 805
 RESULT 18
 Q9VJVL PRELIMINARY; PRT; 792 AA.
 AC Q9VJVL 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DR 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE CG17988 protein.
 GN ANCE-3 OR CG17988
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidae; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; Pubmed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brannon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bens P.V., Berman J.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Bottler P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasner K.,
 RA Glodek A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshirei A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson W., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003641; AAF53356.1; -
 DR FlyBase; FBgn0032536; Ance-3.
 DR InterPro; IPR001548; Peptidase_M2.
 DR InterPro; IPR00130; zn_MPeptidase.
 DR Pfam; PF01401; Peptidase_M2; 1.
 DR PRINTS; PR00791; Peptidase_M2.
 DR Prodom; PD004184; Peptidase_M2; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SO SEQUENCE 792 AA; 92321 MW; 757EC472089022B0 CRC64;

Query Match 19.6%; Score 841; DB 5; Length 792;
 Best Local Similarity 30.5%; Pred. No. 7,8e-53;
 Matches 196; Conservative 105; Mismatches 224; Indels 118; Gaps 14;

QY 48 WNYNTNTEENVQNMNNAAGKWSAFLKEQS--TLAQMT-----PLQEQNTYK 94
 DB 158 YNNNNPNVVEFV-GINNRRFENPFLSNODRFNLQGYLERORYQDRRYQDELEKRLIL 216
 QY 95 L-----OLQALQNGSSV 107
 DB 217 LVESDQKSELECTANLNAOQRYVEFORITAEQSKRIKDLFDRLRLTQMLQSEVGPNA 276
 QY 108 LSEDKSKRLNTILNTMSTIYTGKVCNPDNPOEC-LILEPGLNEIMANSIDYNERLWME 166
 DB 277 LPLDVBRYNRLNLEMLFVNSAETICAYQOPQCDLHYITQLKIDYAKSDMDELQMTW 336
 QY 167 WSRSEVGKOLRPLYEEVYVLKNEBARNHEDYDYGKDEYVNGVDGYTSRQGLTEY 226
 DB 337 EYHKKAGGMDSEYQOLIDVQAEVAYVNNVTNGEYWAYE-----SGNFRQ 386
 QY 227 EHTFEETKPLYEHLHAYVRKIMAY--PSYISPIGCPAHLIDGMGRFNTNLSLV 285
 DB 387 DIWEOILPYLEGLHAYVRKLDYGPDRINRIAPISHILGMVGSNSVLDLIT 446
 QY 286 GQKNIDVTDAVDAQDAQRIEKAKEKFEVSVGLPMMTOGFWMENSLTPGVQKAVCH 345
 DB 447 PGRLIDVTPEVNGVGTPLQMLADEEFTSINMSAVGEFRNSLFEQPLD-RVYLCE 505
 QY 346 PTANDL-GKGFRLMCTKYTMDFLTANHEMGIQDYAMAYAOPELLRNGANGEGHEAV 404
 DB 506 PSAMDFCNHRHDFRVKICTDINQSRSLISVHEMAHIQYFLQYRHLPKIFRNGANPAFRQAV 565

```

OY 405 GEIMSLATPRHLKSLGSLSPDFEONETEINFLKQALITVGLPTMYLKKRMVNF 464
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 566 GDAIGLSVSTRPHIOTLRLORSIDE--SYDINTLFTMAIDKAVLPALSLDWMRDVDF 624
OY 465 KGEIPKDDMKRMKREIVEVPEVPHDETYCDPASLFHVSNDYSFIRYTYTLTQFO 524
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 625 SGNARKRTMNCMYNN-----LRFESTVLOFO 651
OY 525 FOEALCOQAKH-----EGPLHKCDISNSTAGCKLFNMLRKGSEPTTALENVGAKN 578
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 652 IYRGICRSRGYVPGDPRKPLHCDIYRQPAAGNMLTKIMSKGASQFOWEVLTELRGR 711
OY 579 MNVPLIFEPFLTWLKDQON--KNSFVGWSTDMSPYADQSI 619
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 712 LDGTALREYFAPLEBMLRQENLRTNEYGVGWYD-GDICKRSIE 753

RESULT 19
OY 018581 PRELIMINARY; PRT; 907 AA.
AC 018581;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 101.1 kDa protein.
GN C42D8.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxID=6339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX Hallsworth K.;
RT "The sequence of C. elegans cosmid C42D8.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: U56966; AAA98719.1; -.
DR InterPro: IPR001548; Peptidase_M2.
DR Pfam: PF01401; Peptidase_M2; 1.
DR ProDom: PD004184; Peptidase_M2; 1.
KW Hypothetical protein.
SQ SEQUENCE 907 AA; 101086 MW; F9B8B5C0F9BC5AA3 CRC64;

Query Match 15.0%; Score 642.5; DB 5; Length 907;
Best Local Similarity 27.0%; Pred. No. 3.1e-38;
Matches 172; Conservative 136; Mismatches 289; Indels 41; Gaps 19;
OY 2 SSSSMLLSLVAVTAOSTIEQAKTFLDFNHEADLFYSSSLASWNTNTTEENVON 61
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 SSNWTKNDNLQAPGSIKD--EKIRSLAGYEAEAIKVLKFAVLSGRYRNDKSPSLKLA 217
OY 62 MNNAAGDWSAFLEKOSTLAOMYPLQETIQLTVKIQLOALOQNGSSSVLSDEKSKRLMTIIN 121
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 218 LDEANENLTFEVRSTSQAKQFDMASVTDEKVRKQLGYVSFEQMSALAPSRFADYSQAQA 277
OY 122 TMSIYSTGVKCNPDNQECILLEPGNEMANSLOYNEMLMWESRSRVGQLRPLYE 181
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 278 ALNDSKDSITCDKDVPPPCALQKIDMSIFRNEKDKASRLQHLWVSIVYTAIAKS-KPSYN 336

```

```

OY 182 EYVVLKEMARAHNHYEDYCYHMGDYEVNG-VDCYIYSRQGLIEDVHHTPEIKPLYEH 240
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 337 NITTSNEGAKLNGFANGGAMRSAPDMSKVKHAEF---DLNKQIDKITYSTIQPYQLL 393
OY 241 HAVVRKAKIMAY--PYSISPGLPAHLIGDMGRWTNLSLTYPEGQKPNIDYDANV 298
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 394 HAYMRQLAGIYNPNVGLSKDGPPIPAHLFGSLDGGWMSAHYEDQTKPEEES--ETPEAML 451
OY 299 D----QAMPQRIEKEAEKFEVSVGLPNTOGFENSMITDQGNQKAVCHP-TAMD-LG 352
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 452 SAVNTQWYTTKMKMEVATAYRFKSGAPPHLPKSVWTSIARWMS-KDMCHPAALADMA 510
OY 353 KGPRLIMCTKYVMDFLANHEMGHIQYDMAAOPFLIRNGANGFEHVAEGLTSLA 412
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 511 PNDFRVACAGQIGEPDFEQASHLLVOTYYOYLTKDQSLFRDQSPVITDAIANFAHLS 570
OY 413 ATPKHLKSLGSLSPDFEONETE-INFLKQALITVGLPTMYLKKRMVNFKEGIPKD 471
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 571 TNPHYLSQKLVSEHLIDKDSYITINKLYKESLESTFKLPFLIADNMKYLEFDGTVPKN 630
OY 472 QMKKMKMKREIVEVPEVPHDETYCDPASLFH--VSNDYSFIRYTYTL-----YQFOF 525
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 631 KLNDRMWELRNKKEGVRSPOPYWTSMLD--ALIHNSVQVHS--PATRTLSYVLKFOI 685
OY 526 QEALCOQAA--KHGFLHKCDISNSTEAGOKLFNMLRLKSEPTTALENVGAKNMVNR 582
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 686 LKLCORELFWISEG---CLISEDTT--EKRETKLSSITWKLALMISGKGLDQ 739
OY 583 PLNVEFPLTWLKDQON--NSFVGWSTDMSPYADQSI 618
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 740 PLEIYEPLINLARNNEIDQYVVGWDSGETPYVEI 777

```

```

RESULT 20
OY 09V520 PRELIMINARY; PRT; 661 AA.
AC 09V520;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE CG8196 protein.
GN ANCE-4 OR CG8196.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Adyaani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman I.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```



```

Db      61  NMNNAADKWSAFLEQSTTAQMPLQEIQNLVFKLQALQOQSSVLSDEKSKRLNTLL 120
QY      121  NTWSTYSGKVCNPDPNPOCELLLEGLNEIMANSIDYERLWAMESRSSEVGKOLRPLY 180
Db      121  NTWSTYSGKVCNPDPNPOCELLLEGLNEIMANSIDYERLWAMESRSSEVGKOLRPLY 180
QY      181  EEEVVLKNEKMARANHYEDYDGYWNGDYVNGDYSGQLIEDVEHFEETIKPLYEHL 240
Db      181  EEEVVLKNEKMARANHYEDYDGYWNGDYVNGDYSGQLIEDVEHFEETIKPLYEHL 240
QY      241  HAYVRKLMNAIPSYISPGCLPAHLGDMGREFWNTYSLTFPGQKPNIDVTAMVQ 300
Db      241  HAYVRKLMNAIPSYISPGCLPAHLGDMGREFWNTYSLTFPGQKPNIDVTAMVQ 300
QY      301  AMDAORIFKEAEKFEFVSGLPMMTQGFENSMULTDPGNQKAVCHPTAMDLSKGFRLM 360
Db      301  AMDAORIFKEAEKFEFVSGLPMMTQGFENSMULTDPGNQKAVCHPTAMDLSKGFRLM 360
QY      361  CTKVMTMDFLTAHHEMGHIQYDMAVAAPFLRLNGANGFHEAVGEIMSLAATPKHLKS 420
Db      361  CTKVMTMDFLTAHHEMGHIQYDMAVAAPFLRLNGANGFHEAVGEIMSLAATPKHLKS 420
QY      421  IGLLSPDQEDNETEINFLKQALTYVGLPTTYMLEKRWMTKEGELPKDQMKKWMEM 480
Db      421  IGLLSPDQEDNETEINFLKQALTYVGLPTTYMLEKRWMTKEGELPKDQMKKWMEM 480
QY      481  KREIIVGVEPVPHDETYCDPASLFHVSNDYSFIRYRTLYLOFOQOALCOAAHHEGLH 540
Db      481  KREIIVGVEPVPHDETYCDPASLFHVSNDYSFIRYRTLYLOFOQOALCOAAHHEGLH 540
QY      541  KCDISNSTEAGOKLFNMLRLGKSEFWTLAENNVGAKNMNVRPLNTEPELFTWLKDQNK 600
Db      541  KCDISNSTEAGOKLFNMLRLGKSEFWTLAENNVGAKNMNVRPLNTEPELFTWLKDQNK 600
QY      601  NSFVGMSTWSPYADOSTIKVRLSKSALGKAYEMNEMNEMLEFSSVAYAMROYFLVYKN 660
Db      601  NSFVGMSTWSPYADOSTIKVRLSKSALGKAYEMNEMNEMLEFSSVAYAMROYFLVYKN 660
QY      661  QMILFGEEDVAVANLKPRISFNFVTAAPKNVSDIIPTEVEKATIRMSKRINDAFRLND 720
Db      661  QMILFGEEDVAVANLKPRISFNFVTAAPKNVSDIIPTEVEKATIRMSKRINDAFRLND 720
QY      721  SLEFLGIQPTLGPNNOPPVSTWMLVFGVGMVYVGVIVILFTGIRDKKKNNARSGENP 780
Db      721  SLEFLGIQPTLGPNNOPPVSTWMLVFGVGMVYVGVIVILFTGIRDKKKNNARSGENP 780
QY      781  YASIDISKENNPGFQNTDDVQTSF 805
Db      781  YASIDISKENNPGFQNTDDVQTSF 805

```

RESULT 2

```

US-08-481-626-2
Sequence 2, Application US/08481626
Patent No. 5801040
GENERAL INFORMATION:
APPLICANT: Soudrier, Florent
APPLICANT: Alhenc-Gelas, Francois
APPLICANT: Hubert, Christine
APPLICANT: Corvol, Pierre
TITLE OF INVENTION: Nucleic Acid Coding for the Human
TITLE OF INVENTION: Testicular Angiotensin Converting Enzyme (ACE) and its
TITLE OF INVENTION: Uses, Especially for the In Vitro Screening for this
TITLE OF INVENTION: Enzyme in the Organism
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,626
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/656,183
FILING DATE: 04-MAR-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 89-09062
FILING DATE: 05-JUL-1989
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 04958-0006-02000
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEO ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 732 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-626-2

```

Query Match 31.3%; Score 1344; DB 1; Length 732;

Best Local Similarity 41.8%; Pred. No. 4,8e-120;

Matches 259; Conservative 119; Mismatches 204; Indels 38; Gaps 10;

```

QY      15  TAAQSTT---TTEQAKTFPKFHEAEDLPYQSSLASWNYNTNTEE-----NQNM 62
Db      61  TSNQSPNLTVDKAEAKFVEEDRTSQVYWNELAEAMNYNTNTEESKILLQNMKQIA 120
QY      63  NNAQDKWSAFLEQSTTAQMPLQEIQNLVFKLQALQOQSSVLSDEKSKRLNTIINT 122
Db      121  NHT-----LKYGTOAKKFDVNOLONTTIKRIIKYQDDERALLPQOEEYKILLD 172
QY      123  MSTYSGKVCNPDPNPOCELLLEGLNEIMANSIDYERLWAMESRSSEVGKOLRPLYEE 182
Db      173  METTYSVATVCHPNC--SCLQLEPDLTNVMASTRKYEDLLWAMEGRDKAGRAILOFYPK 230
QY      183  YVVLKNEKMARANHYEDYDGYWNGDYVNGDYSGQLIEDVEHFEETIKPLYEHLHA 242
Db      231  YVELIQARLNGYVDAGDSWMSMETPSLE-----QDLERLFQELQPLYLNLHA 280
QY      243  YVRKLMNAIPSYISPGCLPAHLGDMGREFWNTYSLTFPGQKPNIDVTAMVQ 301
Db      281  YVRALHRRHYGAQHINLEPRIPAHLLGNMMAQTWNSINIVLVPPSPASMDTTEAMLKQ 340
QY      303  WDAORIFKEAEKFEFVSGLPMMTQGFENSMULTDPGNQKAVCHPTAMDLSKGFRLM 360
Db      341  WTPRRMFKKADDFETSLGLPVPPEFWKSMLEKFTDREVVCHASADFYNGKDFRIKQ 400
QY      361  CTKVMTMDFLTAHHEMGHIQYDMAVAAPFLRLNGANGFHEAVGEIMSLAATPKHLKS 420
Db      401  CTTVMLEDLVVAHHEMGHIQYFMQKDPVALREGANGFHEALGDVALSVSTPKHLHS 460
QY      421  IGLLSPDQEDNETEINFLKQALTYVGLPTTYMLEKRWMTKEGELPKDQMKKWMEM 480
Db      461  IMLLSSEGGSD-EHDIINFLMKALDKIAFIPFSYLVQMRKRVDPGSIITKENYQEWMSL 519
QY      481  KREIIVGVEPVPHDETYCDPASLFHVSNDYSFIRYRTLYLOFOQOALCOAAHHEGLH 540
Db      520  RLKYGGLCPVPRYRQGGDDPKAKRHHSVSYIYKFSIIQFOFHALCOAGHNGPRL 579
QY      541  KCDISNSTEAGOKLFNMLRLGKSEFWTLAENNVGAKNMNVRPLNTEPELFTWLKDQNK 600

```

Db 580 KCDIYQSKGAGQRLATAMKLGFSRPMPEAMQILITGPPNNMSASAMLSYFKPLDMLTENE 639
 QY 601 --NSFVGW-STDWSPYADQS 617
 Db 640 LHGEKLGMPQYNMTSPNSARS 659

RESULT 3
 US-08-989-299-4
 ; Sequence 4, Application US/08989299
 ; Patent No. 6194556

GENERAL INFORMATION:
 APPLICANT: Acton, Susan L.
 APPLICANT: Robinson, Keith E.
 TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
 TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/989,299
 FILING DATE: 11-DEC-1997
 CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
 NAME: Arnold E., Beth
 REGISTRATION NUMBER: 35,430
 REFERENCE/DOCKET NUMBER: MIA-025.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 732 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-989-299-4

Query Match 31.3%; Score 1344; DB 4; Length 732;
 Best Local Similarity 41.8%; Pred. No. 4, 8e-120;
 Matches 259; Conservative 119; Mismatches 204; Indels 38; Gaps 10;

QY 15 TAAOS-----TTEBOAKTFLLDKFNHEADLFFYQSSLSAMWYNTNITE-----NVQNM 62
 Db 61 TSAOSPMLYVDEAEASKVEYEDRTSQVNNNEFAEAMNNNTNITETSKILLQKMQA 120
 QY 63 NNAGDKNSAFLKEQSTLAOMYPLQELQNLVTKLQALQONGSSVLSDEKSKRLNTLNT 122
 Db 121 NHT-----LKYGTQARKFEDVQNLQNTIKRIKKVQDLERAAALPAQELREYNKILD 172
 QY 123 MSTYSTGKVCNPDNPOBCLLEPGLEIMANSIDYNERMAMESRSEVGKQRLARLYEE 182
 Db 173 METYSYATVCHPNNG--SCDLEPDLNNVATSKREDLLMAGWKRAKRAIILQFYRK 230
 QY 183 YVVLKNEARAHNEYEDYRGDYGVGVDYSGQLLEDEVENTFEERIKPLYEHLA 242
 Db 231 YVELINQAKRLNGYDADGSRMSYETPSLE-----QDLERLFOELQPLYLNIHRA 280
 QY 243 YVRATLMNAV--PSYISPIGCLPAHLGLDMGREFNTNLSLTVPPGQKPNIDYDAMDQA 301
 Db 281 YVRALHRYGAQHINLEGPPIPAHLILGNMAQOTSNITDIYLVPPSPASMDTTEAMLKQG 340

QY 302 WDAQRIKFEAKETFFVSGLEPNMTOGEWNSMLTDEGNVOKAVCHPTAMDLAGK--DEFILM 360
 Db 341 WTPRRMFKEADDEFTSTIGLLPPEPEFWNKSMLKEPTDGEVYVCHASANDFYNGKRFRIKQ 400
 QY 361 CTKYTMDFLFAHHEMGHIOYDMAAOPFLLRNGANEGEHEAVGEIMSLSATPKHLKS 420
 Db 401 CTYVNEDELYVAHHEMGHIOYFMQKDLRYALLEGANPGFHEAIGDYALSLVSTPKHLHS 460
 QY 421 IGLSPDQEDNETETINFLKQALTYGTLPTFTMLEKRMVYKEIKPKDMKKMMEM 480
 Db 461 INLSSREGSD--EHDINFLMKMALDKIAFIPFSYLVQDMWRVFDGSTIKENTINDENMSL 519
 QY 481 KREIVGVEEPHPHDETCDPASLEFVNSDYSFTRYTRTLTYQFOFQALCOAKHEGPIH 540
 Db 520 RLKYOGCLPPYRTQGFDEQAKHISVPYIRYFVSFIQGFHEALCOAAGHTGPLH 579
 QY 541 KCDISNSTEAGOKLFNNMLRKGSEPTWLALENVGAKNMNVRLNYEPFLTWLKDQK 600
 Db 580 KCDIYQSKGAGQRLATAMKLGFSRPMPEAMQILITGPPNNMSASAMLSYFKPLDMLTENE 639
 QY 601 --NSFVGW-STDWSPYADQS 617
 Db 640 LHGEKLGMPQYNMTSPNSARS 659

RESULT 4
 US-08-989-299-7
 ; Sequence 7, Application US/08989299
 ; Patent No. 6194556

GENERAL INFORMATION:
 APPLICANT: Acton, Susan L.
 APPLICANT: Robinson, Keith E.
 TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
 TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/989,299
 FILING DATE: 11-DEC-1997
 CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
 NAME: Arnold E., Beth
 REGISTRATION NUMBER: 35,430
 REFERENCE/DOCKET NUMBER: MIA-025.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1306 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein

Query Match 31.2%; Score 1337; DB 4; Length 1306;
 Best Local Similarity 41.7%; Pred. No. 5, 9e-119;
 Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

QY 20 TTEBOAKTFLLDKFNHEADLFFYQSSLSAMWYNTNITE-----NVQNMNAGKQKSA 71

```

Db 644 TDEAFASKFVEEYDRTSYVWNEEYAEAMNNTNTTETSKILLQKNQJANH----- 697
QY 72 FLKEOSTLAQMPLOEIQNLVYKLOLQOOGSSVLSDESKRLNTLNTMSTYSTGK 131
Db 698 --LKGTQARKKDVNOLOTTIKRIKKVQODIERALPAQLEEFNKILLDMETTYSAT 755
QY 132 VCNPNPQECLEELFGLMEINANSIDYERLWAMESSEVQKQRLPYEEYVLKNEMA 191
Db 756 VCHPNG--SCIQLEPDLINVAATSKRYKIDLLWAMEGMDKGRALILQFYPRYVELIQQA 813
QY 192 RANHVEDYDGRDYEVNGVDYDSRGQLEIEVHTFEERIKPLEYHLAAVYAKLMA 251
Db 814 RUNGTVADGDSWRMYETEPSE-----QDLERLFOQLVYLNLAHYRRALHHR 863
QY 252 Y-PSYISPGCLPAHLIGDMWGRFTNLYSLTVFQGRPNIDVTADAMDAQRIEKE 310
Db 864 YGAGHINLEGP1PAHLIGDMWGRFTNLYSLTVFQGRPNIDVTADAMDAQRIEKE 310
QY 311 AEKPFVSGVGLPMTQGEWENSMLTDPGNVQKAVCHPTAMDLGKG-DPRILMCTKYTMDPF 369
Db 924 ADDEFTSLIGLLPVPPEFNNKSMLEKPTDGREVVCHASAMDFYNGKDFRIKQCTVLEDL 983
QY 370 LTAHEMGHIQDMAAQAAPFLIRNGANEGFHEAVGIMSLSAATKHLKSLGSLSPDQ 429
Db 984 VVAHHEMGHIQDMAAQAAPFLIRNGANEGFHEAVGIMSLSAATKHLKSLGSLSPDQ 429
QY 430 EDNETEINFLKQALITVGLPPTVYMLEKRWYFEGELPKQDMKRWEMKREIYVGE 489
Db 1044 SD-SHDINFLKQALITVGLPPTVYMLEKRWYFEGELPKQDMKRWEMKREIYVGE 489
QY 490 PVPHEFYCDPASLFHVSNDYSFIRYTRTLVQFOFQALCOAAKHEGLHCDISNSTE 549
Db 1103 PVPHEFYCDPASLFHVSNDYSFIRYTRTLVQFOFQALCOAAKHEGLHCDISNSTE 549
QY 550 AGQKLFMLRLKSGSEPTLALENVGAKMNVRLINTEPELFTWLKQDNK--NSFVGM- 606
Db 1163 AGQKLFMLRLKSGSEPTLALENVGAKMNVRLINTEPELFTWLKQDNK--NSFVGM- 606
QY 607 STWSPFYADQS 617
Db 1223 QYMTWSPFYADQS 617
Db 1223 QYMTWSPFYADQS 617

```

RESULT 5
 US-08-989-299-5
 ; Sequence 5, Application US/08989299
 ; Patent No. 6194556
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan L.
 ; TITLE OF INVENTION: ANGIOGENIN CONVERTING ENZYME HOMOLOG
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109-2170
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/989,299
 ; FILING DATE: 11-DEC-1997
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Arnold E., Beth
 ; REGISTRATION NUMBER: 35,430

```

; REFERENCE/DOCKET NUMBER: MIA-025.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-989-299-5

Query Match 31.1%; Score 1334; DB 4; Length 732;
Best Local Similarity 42.6%; Pred. No. 4,4e-119;
Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

QY 20 TIEQAKTFLDKFNHEADLFYQSLASWNYNTNTEENVQNNNAGDKWSAFLEKQSTL 79
Db 69 TDEAKAERFVEEYDRTSYVWNEEYAEAMNNTNTTETSKILLQKNQJANH----- 697
QY 80 AQMTPLQEIQNLVYKLOLQOOGSSVLSDESKRLNTLNTMSTYSTGVCVCPNDPQ 139
Db 129 AKTQDVSNFQNSSIKRIKQLQNDRAVLPKPELEYNOILDMETTYSLSNICYTNG-- 186
QY 140 ECLLEPGLMEINANSIDYERLWAMESSEVQKQRLPYEEYVLKNEMARAHYEDY 199
Db 187 TCMPELEDLINVAATSKRYKIDLLWAMEGMDKGRALILQFYPRYVELIQQA 813
QY 200 GDYWRGDYEVNGVDYDSRGQLEIEVHTFEERIKPLEYHLAAVYAKLMAVPS- YISP 258
Db 247 GDSWRSLSDEDNLE-----QDLERLFOQLVYLNLAHYRRALHHRYSGEYINL 296
QY 259 IGLPFAHLIGDMWGRFTNLYSLTVFQGRPNIDVTADAMDAQRIEKEKFEFVS 310
Db 297 DGP1PAHLIGDMWGRFTNLYSLTVFQGRPNIDVTADAMDAQRIEKEKFEFVS 310
QY 319 GLPNTQGEFWEENSMLTDPGNVQKAVCHPTAMDLGKG-DPRILMCTKYTMDPF 369
Db 357 GLPNTQGEFWEENSMLTDPGNVQKAVCHPTAMDLGKG-DPRILMCTKYTMDPF 369
QY 378 HIQDMAAQAAPFLIRNGANEGFHEAVGIMSLSAATKHLKSLGSLSPDQOEINEIN 437
Db 417 HIQDMAAQAAPFLIRNGANEGFHEAVGIMSLSAATKHLKSLGSLSPDQOEINEIN 437
QY 438 FLIKQALITVGLPPTVYMLEKRWYFEGELPKQDMKRWEMKREIYVGEVPHDET 497
Db 476 FLIKQALITVGLPPTVYMLEKRWYFEGELPKQDMKRWEMKREIYVGEVPHDET 497
QY 498 CDPASLFHVSNDYSFIRYTRTLVQFOFQALCOAAKHEGLHCDISNSTEAGQKLF 550
Db 536 CDPASLFHVSNDYSFIRYTRTLVQFOFQALCOAAKHEGLHCDISNSTEAGQKLF 550
QY 558 IRLKSGSEPTLALENVGAKMNVRLINTEPELFTWLKQDNK--NSFVGM-STWSP 612
Db 596 IRLKSGSEPTLALENVGAKMNVRLINTEPELFTWLKQDNK--NSFVGM-STWSP 612

```

RESULT 6
 US-08-989-299-8
 ; Sequence 8, Application US/08989299
 ; Patent No. 6194556
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan L.
 ; TITLE OF INVENTION: ANGIOGENIN CONVERTING ENZYME HOMOLOG
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA

RESULT 7

[illegible]

RESULT 11
 US-09-440-325A-1
 ; Sequence 1, Application US/09440325A
 ; Patent No. 6280994
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Paul O.
 ; TITLE OF INVENTION: Zacl: A Human Metalloenzyme
 ; FILE REFERENCE: 98-79
 ; CURRENT APPLICATION NUMBER: US/09/440.325A
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FASTSEQ for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 694
 ; TYPE: PRP
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)...(694)
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-440-325A-1

Query Match
 Best Local Similarity 23.1%; Score 990; DB 4; Length 694;
 Matches 221; Conservative 104; Mismatches 228; Indels 30; Gaps 15;

22 EEOAKTELDKFNHEAEDLEFYOSSLASWNTNTTEENOVNMMNAGDKMSAFLEKOSTLAQ 81
 53 ETEKTFLOEYDQGEVVLNFKMEATWNTYITNRKNOEMKDMER-SQMIFGTQAH 111
 82 MPLOEIONLVKLOALQNGSSVLSDEKSRKLTNTLNTMTSTGYVCNDPDRQC 141
 112 LEKVQFKDPVNGMLSKLONIDKALSKDELREYNELLAXLETGYMAOVCLNECP-C 169
 142 LLEPGNEIMANSLDYNERLWAMESWSEVQKOLRPLYEYVYVKNEMARANH-YEDYG 200
 170 LLSLESEL-EVMAISRDEKELLMAMOGMODAVGRICITTFEYVLSNKAQOLNGYXKDMG 228
 201 DYRGGYEVNGVDGYDSRGOLIEDVETFEIKPLYEHLNAYRAKLINAY-PSYISPT 259
 229 ALMHSKTESDYLE-----ODLRLOELRPPLYLNPHTYVRALHRYGELLDR 278
 260 GCLPAHLIGD-MMGRFNTNLYSLVPGQKPNIDVTDAVDAQMAORI-FKAEKFEVS 317
 279 GFLPAHLIGENTLAQSWNLIDVLPFLKIPEDVTKIMKYOHMKPEKMLEEAEFFTY 338
 318 VG--LPMNTOGFEMENSLDTPGVCQAVCHPTAMDIGK-GDFRLMCTKYTMDPFLTAH 374
 339 LGIALPAPPSFWKTLMLRPTDGRREVCHISANWFYQODDFRIRKCAEVTEDEPLSIFH 398
 375 EMGIQYDAVYAAOPFLLRNGANEGFHAVEIMSTSAATPKILKSGILSPFOEDNET 434
 399 EMGHFYFLQYKNLSIFRTGANPAFEAVGSYITTSASHKHLNLIGILS--LLED--- 453
 435 EINFELKQALITVGLPTTYMLEKRMWVFKGEIPPDQMMKRWMMKR-ETIVGVEEVPH 493
 454 EVNFMHLALKEIAFIPEGYMLDFRKVFVDGTMKDIYNQWMLRLRLKQIGCPAIPH 513
 494 DETYDPAISLHVANDYSFT-RYTYRTLYVFOFOBALCOAKHSGPLAKDCDISNSTEAG 552
 514 SEEDFDPKAFHPSAGVPIYTRRFTSLVLOFHFETLCKASGHGPHQCDIYNKSIACK 573
 553 KLFNMLRIGKSEPMITALENNVGAKNMVRLPLNVEPELFTWL 595
 574 LL--ALKIGSSKFWPEVLKMLTGESEVSTNVEKTYENKPLLTWL 614

RESULT 12
 US-08-989-299-12
 ; Sequence 12, Application US/08989299
 ; Patent No. 6194556

GENERAL INFORMATION:
 APPLICANT: Acton, Susan L.
 APPLICANT: Robinson, Keith E.
 TITLE OF INVENTION: ANGIOGENESIS CONVERTING ENZYME HOMOLOGY
 NUMBER OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
 CORRESPONDENCE ADDRESSES: 14
 ADDRESSES: FOLLEY, HOAG & ELLIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/989,299
 FILING DATE: 11-DEC-1997
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Arnold E., Beth
 REGISTRATION NUMBER: 35,430
 REFERENCE/DOCKET NUMBER: MIA-025.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 907 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-989-299-12

Query Match
 Best Local Similarity 15.0%; Score 642.5; DB 4; Length 907;
 Matches 172; Conservative 136; Mismatches 289; Indels 41; Gaps 19;

2 SSSSWLLSLVAVTAOSTIEEOAKTFLDKFNHEAEDLEFYOSSLASWNTNTTEENOVN 61
 160 SSNMYKTNLOAPGISIKD--EKLRSMLAGTEAEKIKYLRVALSGRNYRNDSPSKIA 217
 62 MNAGDKMSAFLEKOSTLAQSWNLIDVLPFLKIPEDVTKIMKYOHMKPEKMLEEAEFFTY 338
 218 LDEAEVNLTFEVSSTMSQAKQFDMASVDEKVMRQLGVSEFGMSALATSRPADYQA 277
 122 TWSTIYSGVVCNDPDRQCILLEPGNEIMANSLDYNERLWAMESWSEVQKOLRPLYE 317
 278 ALNRDSKSTIDKQVPPCALQKIDMSIFRNEKASRLQHLWASYVAIAKS-KPSYN 336
 182 EYVVLNEMARANHEDYGDYRGDEYVNG-VQGYDSRGOLIEDVETFEIKPLYEHLNAY 493
 337 NITTIENEGAKLNGFANGAMRBSAFDMSSKVAKAE---DLNKQIDRYSTLOPFTQL 513
 241 HAYVRAKLINAY-PSYISPTGCLPAHLIGDMKGRFNTNLYSLVPGQKPNIDVTDAV 552
 394 HAYMRQLAGIYSNPVGLSKDGPFAHLFGSLDGDMSAHYEQKPFEEES--ETFEAAL 595
 299 D---QAMDQORIFKAEKFEVSVGLPMNTOGFEMENSLDTPGVCQAVCHP-TAMD-LG 352
 452 SAFNTQNYTTKKMFYAIYRFRKSAGFPPLPKSYWTSSTIFARWVS-KDMICHPAALDMRA 510
 353 KGFRLMCKRYTMDPFLTAHHEMGHIQYDAVYAAOPFLLRNGANEGFHAVEIMSTSAAT 513
 511 PNDFRKACAQOUGEDEFOASHLVQTYOYLTKDOSLLFRQASPVITDAIANAFHLS 570
 413 ATPKHLKSGILSPFOEDNETE-INFLKQALITVGLPTTYMLEKRMWVFKGEIPPDQMMK 595
 571 TNPHTLYSOKLVPESEHLDIKDSYITNKLKRESLSEFTKLPITADAMKRIELPDTGTPKN 630

us-09-978-385-2.raii

[illegible]

US-06-043-193B-15
Sequence 15, Application US/08645193B
Patent No. 5962253
GENERAL INFORMATION:
APPLICANT: Kupke, Thomas
APPLICANT: Goltz, Friedrich
APPLICANT: Kempter, Christoph
APPLICANT: Jung, Gunther
TITLE OF INVENTION: Oxidative Decarboxylation of Peptide
TITLE OF INVENTION: Catalyzed by Flavoprotein EpiD
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005

```

? COMPUTER READABLE FORM.
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/645,193B
? FILING DATE: 13-MAY-1996
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Esmond, Robert W.
? REGISTRATION NUMBER: 32, 893
? REFERENCE/DOCKET NUMBER: 0652.1540000
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 371-2600
? TELEFAX: (202) 371-2540
? INFORMATION FOR SEQ ID NO: 15:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 990 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: not relevant
? MOLECULE TYPE: protein
?
? US-08-645-193B-15

```

Query Match	2.98;	Score 124;	DB 2;	Length 990;
Best Local Similarity	18.28;	Pred. No. 0.013;		

Matches	153;	Conservative	121;	Mismatches	257;	Indels	310;	Gaps	40;
---------	------	--------------	------	------------	------	--------	------	------	-----

```
07 I M S S S W L L S V A V T A O S T I E Q A K F L D R N H E -----AEDL--FYÖSSLASWNY 50  
   :|::: | || ::||: |  
08 :|::: | || ::||: |
```

324 ISEAAVILMLLSPNHGCTKTIRNYHEFFEMDKYGFQVLNKLQLLSDINGFGYPPKDSYSF 383

```

51 NNTTEENVQNMNNAGDKWSAFLEQOSTLAQMYPLEIQNLTVKIQLOALQONSSVLSL 110
: || |||||:

```

```

384 SNNI-----AFLEK-----YLAIQNNSHIETE 408

```

```
III DSKRLNTILNTMSTIYSTGKVCNPDNPQECLELLEPGLNEIMANSLDYNERLWAMESWRS 170
```

409 NDVKNLEK-NNTYSKINA-----PVSTEL-----YSEIYF-----437

```

01      QY      17 EVGRKQRLPELVEEVLVKNENAKRANHEDVDGYRDGYEVNVDGYDYRSNGOLIEDEHTF 230
02      Db      438 ---GNSIKG-YEDFAVNISPILIGSFNAGATGRF-TGNFNIR-----KKNOLOKEIYA-- 448
03      QY      231 EIKPFLYEHLHAYVRAKIMAAVSYSISPIGCLPPAHLHGDMRGFTNLXSLTVPGCGKP 290
04      Db      485 -----HYNNYNENNDELFEISQLNAPLRINSRNVNLNN--NRINYTCNLNLP---KSD 531
05      QY      291 IDVTDAWDAQMAOQRERK---AEFEYSVGLPMNTOG-----FWENSMLTDPG 337
06      Db      532 IDINDIFIGATFKNLIVLSKHDSRIIVFSNSMFYEGSELYKFLREISPEKTKFIQPI 591
07      QY      338 N-----VQKAVCHPTAMDLEGDFRILMCTKTMTMODPLAHNHGHCIQYDK 383
08      Db      592 TEECIDLSPCPRIIIYKNIILPAWKINSEMFE---TEWMLNRFATI-REKMHIIPKV 647
09      QY      384 AYA-----AAPFLIRGANGCFHEAVGETMS----- 409
10      Db      648 IIAFGDWRLLNLNDNKHLILIKELKKGHRITLESFINESNNERMLEIVTPLYKTSL 707
11      QY      410 --LSAATPK-----HLK--SIGLSPDPOEDNETEIINFILKOALTIYGLPPTYME 457
12      Db      708 KEOSEIIPKRNRKHPNNLKDWESIHSIPKYQD----NFIDYLLPITTELKANVF 762
13      QY      458 KMFMVAFKEEIKPKDOMKMKWMKREIYGVEYPVPHETYCDDASLFHWSDSYTRY- 516
14      Db      763 KFFYIKFKND--EDFIK-LRLURE-----DEDY-----SOJYSFKMKK 798
15      QY      517 ----TRITYQOFQDALCOAAKHGGPLHKDISNSTEAGOKLFNMLRGKSEPWTIAL 571
16      Db      799 DYCLNSELYDYSIVDYBEVYRYGPHVEDIENFPYDOSIL--SINIQSE-FKIKE 855
17      QY      572 NAVGAKNNVNPRLNYEPFETLWKDNKNSYFGWSTMSPVADOSIKVRSLSKALGD 631
18      Db      856 FIYA--ISIDEFLDYLE-----INNSEEKLINNA--ED 886
19      QY      632 AEENDNMFYFRSSVAYANROFLKVKKNOMLIGEDEVVANLKPRTISENFVYAPKNV 691
20      Db      887 LYRSD-----IREY-----KNLAKL-----TNKRD 909
21      QY      692 SDIP-----RVEEKAIRMSRRINAF--RLDNSILETFIQ 728
22      Db      910 YEILKEPFNLHEPLFMKISILENLKKTLOKSITSRSKISGFHMRCN---RIGIN 965
23      QY      729 P 729
24      Db      966 P 966
25
26 RESULT 14
27 US-08-392-625-20
28 # Sequence 20, Application US/08392625
29 # Patent No. 5837485
30 # GENERAL INFORMATION:
31 # APPLICANT: Enlian, Karl-Dieter.
32 # APPLICANT: G tz, Friedrich
33 # APPLICANT: Schnell, No. 5837485bert
34 # APPLICANT: Augustin, Johannes
35 # APPLICANT: Engelke, Gerhard
36 # APPLICANT: Rosenstein, Ralf
37 # APPLICANT: Kaletta, Cortina
38 # APPLICANT: Klein, Cora
39 # APPLICANT: Wieland, Bernd
40 # APPLICANT: Kupke, Thomas
41 # APPLICANT: Jung, G nther
42 # APPLICANT: Kellerer, Roland
43 # TITLE OF INVENTION: Biosynthetic Process For The Preparation
44 # TITLE OF INVENTION: Of Chemical Compounds
45 # NUMBER OF SEQUENCES: 42
46 # CORRESPONDENCE ADDRESS:
47 # ADDRESSEE: Sterne, Kessler, Goldstein & Fox
48 # STREET: 1100

```

Fri Mar 14 10:00:48 2003

us-09-978-385-2.ra1

Page 10

CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/392,625
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/876,791
FILING DATE: 30-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0652.0980002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 990 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-392-625-20

Query Match 2.8%; Score 120; DB 2; Length 990;
Best Local Similarity 17.8%; Pred. No. 0.031;
Matches 152; Conservative 121; Mismatches 246; Indels 334; Gaps 41;

QY 1 MSSSWLLSLVAVTAOSTIEQAKTPIEDKNEH-----AEDL--FYOSSIASMNY 50
DB 324 ISEAAYILMLSPHEFGTITIRNYHEEFMDKGFEDQLYNKQLSDINGFPYPRKDSYF 383
QY 51 NTNTEENVQNMNAGDKSAFLKEOSTLAQWPIQEIQNTLVKIQALQOQSSVLS 110
DB 384 SNNI-----AFLEK-----YLLAIQNNHIEITTE 408
QY 111 DKSRRLTILNTMSTIYSGKVCNPNPOECILLEGKLEIMANSLDNERLMAESWRS 170
DB 409 NDVKRLK-NNTVSKINA-----PVSTET-----YSEIIF----- 437
QY 171 EVGKQRLPEYEEYVLKKNEMARAHYEDYGDYWRGDYEVNGVDYDYSRGQLIEDVETP 230
DB 438 --GNSIKG-YEDFAVISPIIGSFAGATFGRP-TGNFNK-----KKNQLOKEIVH-- 484
QY 231 BEIKPLEYELHAYVRAKIMNAYPSYISPIGCLPAHLLGDMWGFMTNLISLVPPGOKPN 290
DB 485 -----HNNYMNENDELISQINERAPLNSRVNLIIN--NRITYTCLNLMP--KSD 531
QY 291 IDYEDAVDAQMDAQRPKF--AEKFEVSVGLPNNMTOG-----FWNSMLTDPG 337
DB 532 IDINDIFGATFNKLIYISEKHDSRIYFVSNSMNYEFGSELYKFLREISFEKTRFOP 591
QY 338 N-----VQAVCHPTAWDLG-----G 354
DB 592 TEGIDSLPFCRIYIKNIIILPRTWKINSEMESETENMLNRRATIRKWHIIPDYIANG 651
QY 355 DFIILMCTVMTDDELTAHH-----EMGHIOYDMAYAAOPELLRNGANGEHAYG 405
DB 652 DNLLL-----NLLNKHLLIILKELKHGRIR-----IIESEFINESNNERML 694
QY 406 EIMS-----ISAATPK-----HLK-----SIGLSPPROEDMETELNFLQALT 445
DB 695 EIVTPLYKKTSLEKQSFIIIPKRNKHFNNLKDWFESIHLSIKTYQD-----NFLQDYLLP 749
QY 446 IVGTLPTTYLKKRWMMVKGELPKDOMMKKWEMKREIVGVVPEVPHDETCDPASLFH 505
DB 750 FTELKVNNEFNKFFYIKRED--EDFIK--LRLLRE-----DEDY----- 786

QY 506 VSNDSYFIRRY-----TRLLYQFOPEALCOQAKHEGPHLHKCDISNTEAGOKLENNLR 559
DB 787 -SQIYSFIKNWKDYCLLNSLEYDYSIDYVEYRYRGGHVIEDIENFP-MYDSILSDSIN 844
QY 560 LCKSEPTWLALENVYGAKNMVRPLNYPPLFTWIKDONKNSFVGWSTWSPYDOSIK 619
DB 845 ILQSE-FKIPKEFYIA--ISIDPLDYLE-----INSEK 876
QY 620 VRISLSALGKAYEMNDENEMYLFRRSSVAYAMQYFLAKYKNOMILLEGEDVAVANLKPRI 679
DB 877 EELLNNA--EDLYRSND-----IREY-----KNLLAKL 903
QY 680 SENFETVAKNSDIIIP-----RFEVKAIRMSRINDAF--R 716
DB 904 -----TNPKNDYELIKKEFPNLHEFLFNKISILENKTLQSLVTSRSRIIGSFIMHR 957
QY 717 LNDNSLEFLGIOP 729
DB 958 CN-----RIFGINP 966

RESULT 15

US-08-466-961A-20
Sequence 20, Application US/08466961A
Patent No. 5843709
GENERAL INFORMATION:
APPLICANT: Etlian, Karl-Dieter
APPLICANT: G tz, Friedrich
APPLICANT: Schnell, No. 5843709bert
APPLICANT: Augustin, Johannes
APPLICANT: Engelke, Gernar
APPLICANT: Rosensteln, Ralf
APPLICANT: Kaletta, Corlina
APPLICANT: Klein, Cora
APPLICANT: Wieland, Bernd
APPLICANT: Kupke, Thomas
APPLICANT: Jung, G nther
APPLICANT: Kellner, Roland
TITLE OF INVENTION: Biosynthetic Process for the Preparation of
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, NW
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,961A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/392,625
FILING DATE: 22-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/876,791
FILING DATE: 30-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/784,234
FILING DATE: 31-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0652.0980004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 990 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-466-961A-20

Query Match
 Best Local Similarity 17.8%; Score 120; DB 2; Length 990;
 Matches 152; Conservative 121; Mismatches 246; Indels 334; Gaps 41;

```

QY 1 MSSSSWLLSLVAVYAAQSTEEBAKTFELKFNH-----AEDL--FYQSSLASMY 50
DB 324 ISEAAYIILMLSPHEFGKTRIRYHEFMDKGFQEDLVMLKQLLSDINGFGCPKDDISF 383
QY 51 NTNTEENVQNMNAGKWSAFLKESQSTLAQWPIQEQNLVKIQLDALQNGSSVLS 110
DB 384 SNNI-----AFLEK-----YLLAIQNNSHIETE 408
QY 111 DKSRLTIITMTSTYSGKVCNPDNPOCLLEPGINEIMANSIDYNERIAMESMRS 170
DB 409 NDVKNLEK--NTVSKINA-----PVSTEL-----YSELYF----- 437
QY 171 EVCKOARPLVEEYVLKEMARANYEDYGDYWRGDEYVNGVDYDSRQGLIEDVHTF 230
DB 438 --GNSIKG--YEDFAVISPLIGSFNAGATFGRP--TGNFNK-----KKQLQKEIVH-- 484
QY 231 EEIKPLYEHLAAVYRAKLNNAPYSISIGCLPAHLGLDMGWFNTNLSLVTPFGOKN 290
DB 485 -----HYNNMENDLEISQLEAPLNSRVNINLN--NRIYTCNLMLP---KSD 531
QY 291 IYVTAMDQADQARIFKE--AEKFFVSGLPNMTG-----FWNSMLTDPG 531
DB 532 IDINDIFGAFNKLYLYSEKHDSRIYVSNMFNFESSELYKFLREISEKTFIQPI 591
QY 338 N-----VOKAUCHPTAMDGR-----G 354
DB 592 TEEGIDSLPFCPRITLYKNIILKPAIWKINSEMESETEMLNFRATIRKWHIPKVIITAF 651
QY 355 DFRIMCTKTYMDFLTAH-----EMGHIOYDAVAAQPPILLANGNEGFHVA 405
DB 652 DNRILL-----NLNDKHLILKELKELKHGRIR-----LLESFINESNMERM 694
QY 406 EIMS-----LSAATPK-----HLK-----SIGLSPDQEDNETEINFLKQALY 445
DB 695 EIVTPLYKTSLEKQSFIIIPKRNKHFNNLKDMFSIHLSIKTYQD-----NFIQDYLIP 749
QY 446 IYGLPPTYLKEMRMWFKGEIPKQDMKKWMEKREIYGVYEPVPHDTYCDPASLPH 505
DB 750 FIEELAVNNFINKFYIKKED--EDFIK--LRLALE-----DEDY----- 786
QY 506 VSNDSYFIRY-----TRTLYQFOFOALCOAAKHGGLHKCDISNSTEAGQKLFNMR 559
DB 787 --SOIYSFINKMKDCLLSELKDYIYDYVPEYRYRGGVHIEDINFE--MDSLLDSIN 844
QY 560 LKSESENNLLENVYGAKMNAVRLNYFEPFLTWLKQDKNSFWQSTWSPVADQSIK 619
DB 845 IIGSE--FRIPEFIVA-----ISIDFLDYLE-----INKSEK 876
QY 620 VAIISKALGKAYEMDNENEYVLRSSVAVAMQYFLAKVKNOMILFGEDEVAVANLKRI 679
DB 877 EEILNNA--EDLYRSND-----IREY-----KNLAKL 903
QY 680 SFNEFVAPKPVSDIIP-----RTEVEKAIKMSRSINAF---R 716
DB 904 -----INPKNDEILKKEFPNLHEFLFNKISILENLKKTLOKSLYTSRSLIGSFTHMR 957
QY 717 LNDNSLEFLGIOP 729
DB 958 CN-----RIFGINP 966
  
```

RESULT 16

US-08-630-916A-46

Sequence 46; Application US/08630916A

Patent No. 6011137

GENERAL INFORMATION:

APPLICANT: Pirozzi, Gregorio

APPLICANT: Kay, Brian K.

APPLICANT: Fowles, Dana M.

TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL

NUMBER OF SEQUENCES: 124

CORRESPONDENCE ADDRESS:

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630,916A

FILING DATE: 03-Apr-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MISROCK, S. LESLIE

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-203

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 896-8864/9741

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:

LENGTH: 683 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-630-916A-46

Query Match

Best Local Similarity 18.0%; Score 116.5; DB 3; Length 683;

Matches 142; Conservative 110; Mismatches 268; Indels 269; Gaps 36;

```

QY 12 VAVTAQSTIEQAKTFIDKFNHAEDELFOSSIASWNTNTNTEENVQNMNAGDKMSA 71
DB 44 LASEPADTVNGESSFAPTN-----ASVTGTPVYSEENMLSPCT----- 85
QY 72 FLKEQSTLAQWPIQET-----QNLVTKIQLDALQNGSSVLSSEKSKRLTIITMTSTI 126
DB 86 -----STVEDPVOAELITSSENNCEIPSTSALESEARSILIEPTNS-----SRSSA 134
QY 127 YSTGVCNPDNPOCLLEPGINEIMANS-----LDYNERIAMES 167
DB 135 FEAASKRQPDGCMQPVROQSG--NANTETLPSCWEQRKDPHGRTYVYVHNRTTWEN 190
QY 168 WRSEVGKOLRPLIEEYVLKEMARANYEDYGDYWRGDE--VNGVGYGYDSRQGLIEDV 226
DB 191 -----POPLPPEGWERVDDRRRYVYVHNRTTWENRPTMESVAVNFQWOSORNOI----- 241
QY 227 EHFEIKPLYEHLAAVYRAKLNNAPYSISIGCLPAHLGLDMGWFNTNLSLVTPFG 286
DB 242 QGAMQOFNORYLYSASMLAE-----NDPYGLPP-----G 272
QY 287 OKPNIDVTDAM-----VDQAMDAQRIKFEAKFEVSGLP 321
DB 273 WEKRVDSITRVYFVNNHNTKTQWEDPRTQGLONEEPLPEGWEI--RTYRGVRYFVDHNT 331
QY 322 NMTQGFWNSMLTDPGNVOKAVCH---PTAMDIGKDDFRLKCTKVTYMDFLTAHENGH 378
  
```

Fri Mar 14 10:00:48 2003

us-09-978-385-2.ral

Db 332 TTT-----FKDPRNGKSSVTGKGPQJAYERG---FR-----WKLAH 364
QY 379 IQYDMAYAAQPELLRNGANGCEHVAEGEIMSLSAATPRHLKSLGLSPDOENETINF 438
Db 365 FRY-----LCOSNMLPSHVK--INVSROTLEDESFQOI-M 396
QY 439 LKQALITVGLPPTYLKRWMMVKRGLPRDQ--MMKKWMEKREIVGVVEP----- 491
Db 397 ALK-----PYDLRRRLVYIFRGEGLDGLGLAREWFLSH--EVLNPRYCLFE 443
QY 492 -FHDETC--DPASLF--HVSNDYSFI-RVYTRLYFOFOEALCOAANHEGPIHK-- 541
Db 444 YAGKNNYCLQINPASTINPDLSY-FCFIGRTIAMLFLFGKFDIGESLPIFKRLSKL 502
QY 542 --CDI-----SNSTEGOKLF--NMRLRGKSEPTLALLENVGAKNMN 580
Db 503 TIKDLISIDTEFYNSLIWIRDNNEEGCEKMTFVSDMKELIKVTSIDLK---GGSNLT 558
QY 581 VRPLWVFEPLFTWLKDKNSFGVSTWSPYADQSIKYRISLKSALDKAYEMNDNM 640
Db 559 V-----TEENKDEYIGLMTW-----RFSKGVQEOTKAFIDGPFNEV 594
QY 641 YLFRSSVAYAMROYFLKVNOMILFGEEDVRVANKPRISFNFVTAAPKNVS---DIIFR 697
Db 595 -----VPLQMLQYFDEKELEWALCGMQEVDLADQWNTVYRHYTRNSKOIIMWQFVKE 648
QY 698 TEVEKAIAM 706
Db 649 TDNEVBMRL 657

RESULT 17
US-08-844-059-2
Sequence 2, Application US/08844059

General Information:
Applicant: Lawlor, Elizabeth
Title of Invention: No. 6001601el Compounds
Number of Sequences: 6
Correspondence Address:
Address: Smithline Beecham Corporation
Street: 709 Swedeland Road
City: King of Prussia
State: PA
Country: USA
ZIP: 19046
Computer Readable Form:
Medium Type: Diskette
Operating System: DOS
Software: FASTSEQ for Windows Version 2.0
Current Application Data:
Application Number: US/08/844,059
Filing Date: 18-APR-1997
Classification: 536
Prior Application Number: 9607999.1
Filing Date: 16-APR-1996
Attorney/Agent Information:
Name: Gimmil, Edward R
Registration Number: 38,891
Reference/DoCKET Number: P31456
Telecommunication Information:
Telephone: 610-270-4478
Telefax: 610-270-5090
Telex:
Information for SEQ ID NO: 2:
Sequence Characteristics:
Length: 665 amino acids
Type: amino acid
Strandedness: single
Topology: linear
Molecule Type: protein

US-08-844-059-2

Query Match 2.6%; Score 113.5; DB 3; Length 665;
Best Local Similarity 19.2%; Pred. No. 0.067;
Matches 135; Conservative 99; Mismatches 231; Indels 239; Gaps 33;

QY 143 LLEPGLNEIMANSLDYNEMLNAMESRSEVQKOLRPLYEEVVLKNNMARANHYEDYDY 202
Db 85 LLDISYDKFRTTDDYHEKVA-----QVERLLAODDIY--LGEY 123
QY 203 WRGDEVNGVDGYDSRGOLIE-----DVEITFEELIPYENLHAYV 244
Db 124 -SGWTSYS--DEEFTTESQLAEVFRDEAGNYGIALPSGHEVWSESTFLRLSKYQR 180
QY 245 RAKIMNAVPSYISYIGLPAHL-----LGMMSGFNTLYSLTVPEGOKPNIDV-TDA 296
Db 181 LVEEFKAPHEITPDGRLENLKNFIEPGLIEDLAASRTFTGVPVPSNPKHVVYVWIDA 240
QY 297 NV-----DOAMDA--QRIEKAKEF-----FVSYGLPN--MTQ 325
Db 241 LINTATLGVADQDEHGNDFKFMNGTVFHWGKDLRHSITVPIILMLMDYKLDRLIAH 300
QY 326 GFV--ENSMLTDGNYGKAVCHPTAMDLGKGFRLMCTKYTMDDELTAHHEKHOYD 382
Db 301 GWFVMDGMSKSKGNVYPRMELVERYGDLRLYRLNLPVGSDFTEPDEDYVGRIVE 360
QY 383 MAYAAQPELLR-----NGANGFHEANGEIMSLSAATPRHLKSLGLSPDOENE 433
Db 361 LANDLGNLNRVYSMINKYFDQIPAYEGVTEPDHVALE--KSLA-----DEFTIME 414
QY 434 TEINELLKQALITVGLT--PPTYLKRWMMVKRGLPRDQ--MMKKWMEKREIVGV 487
Db 415 A-VDY--PRALLEAVWTLISRTNKYIDETAPVYLDKDEALRDLASVSHQASIRVVAHL 471
QY 488 VEPVHDEYCDPASLFHYSNDYSFIRYRTITLYOFOEALCOAANHEGPIHKDISNS 547
Db 472 IEP-----FMETSRAV----- 483
QY 548 TEAQOKLFNMLRLGKSEPTLALENV-----VAKMMNVRPLLYNEPFLFTWLD 597
Db 484 -----LQIGLEE--VSSLENLSLADFPADYVYAKGPIFPRLNMEEL-AIKRE 531
QY 598 QNKSFGVSTWSP-----YAD-QSIKVRIS-----LKSALDKAYEM-- 635
Db 532 QMEGNKPAVEKEMNDEVELKINKDELKFEDEPKVLEIRVAEKEYSKVEGSKLLQRFULD 591
QY 636 -NDNEKYLFRSSVAYAMROYFLKVNOMILFGEEDVRVANKPRISFNFVTAAPKNVSDI 694
Db 592 AGDGEDROLTSGIA---KYV--PNQELVGRKVOIVANKR----- 628
QY 695 IPRTVEKAIAMSRINDAFRL--NDNSLEFLGIOTPLGPPN 735
Db 629 -----KMKKTVSOGMILSHEHGKRLTLTVDPVAV--PN 660

RESULT 18
US-09-431-202-2
Sequence 2, Application US/09431202

General Information:
Applicant: Lawlor, Elizabeth
Title of Invention: No. 6294175el Compounds
Number of Sequences: 6
Correspondence Address:
Address: Smithline Beecham Corporation
Street: 709 Swedeland Road
City: King of Prussia
State: PA
Country: USA
ZIP: 19046
Computer Readable Form:
Medium Type: Diskette
Computer: IBM Compatible

```

OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,202
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/844,059
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gimm, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31456
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-431-202-2

```

```

Query Match      2.6%; Score 113.5; DB 4; Length 665;
Best Local Similarity 19.2%; Pred. No. 0.067;
Matches 135; Conservative 99; Mismatches 231; Indels 239; Gaps 33;

```

```

QY 143 LLEPLNIMANSIDYNERLMAWMSRSEVQKOLRPLYEEVYLKEMARAHYEDYGY 202
DB 85 LLDISYDEFIKTTDYHEKVA-----QVEERLADDDIY--LGEY 123
QY 203 WRGDYEVNGVDYSRGOLIE-----DVEHTTEELKPLYEHAAVY 244
DB 124 -SGWYSVS--DEEFTESSQLAEVPRDAGAVTGTGIAPSGHEVWSESEYFLKSLYODR 180
QY 245 RAKLMAYPSTISPTGLPAHL-----LGMWGRFTNLYSLVYFPGQKPAIDY--TDA 296
DB 181 LVEFFKAHPETIPDGLNEMLRNTEPGLDLAVSRFTFTGVPVPSPKHVVYVWIDA 240
QY 297 MV-----DOAWDA--QRIFKAERF-----FVSGLPN--MTQ 325
DB 241 LLNATATLGAODEHGFEDFNGYVPHMGKDLRFHSITWPIILMLDYLPLRLIAH 300
QY 326 GFV---ENSMILDPGNVAKAACHPTAMDIGKGFRLIMCTKYTMDFLAHEMGIQYD 382
DB 301 GWFYMKDKGKSKSGNVVPEMLVERYGIDPLRYILMRMLPVGSDGTFTPEYVGRINTE 360
QY 383 MAYAAGPFLR-----NGANEGFHAIVGEIMSLSATPKHUKSIGLLSPDQEDNE 433
DB 361 LANDGLNLKRTVSMINKYFDGQIPAYVGVTEFDHVLAEVAE--KSLA---DEHTHME 414
QY 434 TEINFLKQALITVGLT---PFTYMLEKRWMMVFKEGPIKQDM---MKKWMKKEIYGV 487
DB 415 A-VDY---PRALAEVWTLISRTNKYIDETAPWLDKDEALRDQLASVWSHQAISIVVHL 471
QY 488 VEEVPHDEYCDPASLFHYSNDYSFRTYRTRLYQFQEQEALCOAKHGEPLHKDCISNS 547
DB 472 IEP-----FMETSRAY----- 483
QY 548 TEAGORLFLNMLRLGKSEPTLALENV-----VGAKNNVNPRLNTEPFLTWLKD 597
DB 484 -----LTOIGLEE--VSSLENISLADFPADYTVVAKGPIPLRNMEETI--AYTKE 531
QY 598 QAKNSVSGSTWSP-----YAD-QSIRKVIS-----LKSALDKAYEW--- 635
DB 532 QMGKNPRAVEKEMNPDEVELKLINKDEIKFEDDKVEIRHAEVKEVSKYSGSKDLLQFRLD 591
QY 636 -NDNEMYLPRSSVAAYAMROYFLKVNOMILREGEYVRANLKPRTSEFFVTAPKKNVSDI 694
DB 592 AGDGEDROLISGIA---KTY---PNEQELVGKKVQIVANLKPRT----- 628

```

```

QY 695 IPRTVEKAIKMSRSRINDAPRL---NDNSLEFLGICPTGLPPN 735
DB 629 -----KMKKKYVSGGMILSNAHDKLTLITLVDAV--PN 660

```

RESULT 19

```

US-08-491-357-2
Sequence 2, Application US/08491357
Patent No. 5716782
GENERAL INFORMATION:
APPLICANT: Golemis, Erica A.
APPLICANT: Law, Susan
APPLICANT: Estojak, JoAnne
TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING A SIGNAL
TITLE OF INVENTION: MEDIATOR PROTEIN THAT INDUCES CELLULAR MORPHOLOGICAL
TITLE OF INVENTION: ALTERATIONS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dann, Dorfman, Herrell and Skillman
STREET: 1601 Market Street Suite 720
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/491,357
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.
REGISTRATION NUMBER: 36,252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-491-357-2

```

```

Query Match      2.4%; Score 104; DB 1; Length 834;
Best Local Similarity 20.5%; Pred. No. 0.81;
Matches 88; Conservative 59; Mismatches 142; Indels 140; Gaps 19;

```

```

QY 56 EENVQNNMNAAGDKSAFLKE-----OSTLAQWYPLDE--IONITVYLQALQDQSSVLS 109
DB 442 ERHINEIRAVAKVEFLFLEBYLHFKAGVANAACLPBLLHKKMRRELDQRY--EDSHQILS 500
QY 110 E-----DKSRKRLTILNT----- 122
DB 501 QTSHTLNEGSWSLNTILAIKPNKCDLDRFVYVAKTVPDADAKQLTITNTNAEALFRPG 560
QY 123 -----NSTY-----STGKYCNP--DNQEC--LLBPLGLNEIYA---NSLDY 158
DB 561 PGSILHKNPESIMNSTEYPRHGGSGQLHPEDHKAQAHNRKALPPLGSLKEQAPDCSSSDG 620
QY 159 NERLWAME-----SWRSEVGRKOLRPLYEEVYLKNEVARAHYEDYGDYWRGDEYVNGV 212
DB 621 SERSWDDYDYVHLQGEKEFERQCKELKEKIMQKNKMLEHHQ----- 665
QY 213 DGVDYSRQGLIEDVHTTEELKPLYEHAAVYRKLMAVPSYISPTGLCPAHLIGDMWG 272

```

Db 666 ----LSQFOLLQ-----EITKPVENDISKW---KPSQSLPTTNSGVSAODROLLCFYYD 713
 QY 273 RFWTNTXSLTVFPGCKPNIDVTDAMVDQAMQO--RIFKEAEKF-----FVSGGLPNM 323
 Db 714 QCFTHFTSL-----LNALDALFSCVSSAOPPRITVAHSKYVILSAHKLVIIG-DTL 763
 QY 324 TQGFWENSMLTDPGNVQAKVCHPTAWDLGKGFRIIMCTKYTMDF--LTAHHEGHIQY 381
 Db 764 TRQVTAQDITRNKVMSSNQLCEQLK-----TIVMATKMAALHPSTTALQEWVHQVT 815
 QY 382 DMAYAAQPF 390
 Db 816 DLSRNAQLF 824

RESULT 20
 US-08-968-633-2
 ; Sequence 2, Application US/08968633
 ; Patent No. 6100384

GENERAL INFORMATION:

APPLICANT: Golemis, Erica A.

APPLICANT: Law, Susan

APPLICANT: Estojak, Joanne

TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING A SIGNAL

TITLE OF INVENTION: MEDIATOR PROTEIN THAT INDUCES CELLULAR MORPHOLOGICAL

TITLE OF INVENTION: ALTERATIONS

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dann, Dorfman, Herrell and Skillman

STREET: 1601 Market Street Suite 720

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103-2307

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/968,633

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/491,357

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Reed, Janet E.

REGISTRATION NUMBER: 36,252

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 563-4100

TELEFAX: (215) 563-4044

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 834 amino acids

TYPE: amino acid

STRANDEDNESS: NO, 6100384 Relevant

TOPOLOGY: NO, 6100384 Relevant

MOLECULE TYPE: Protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-968-633-2

Query Match 2.4%; Score 104; DB 3; Length 834;

Best Local Similarity 20.5%; Pred. No. 0.81;

Matches 88; Conservative 59; Mismatches 142; Indels 140; Gaps 19;

QY 56 EENVYOMNNAAGDKWSAFLEK-----QSTLAQMYPLQE-TONTLVKLOLQALQNGSSVLS 109

Db 442 EKHINEIRYAVDKVELFLKEYLHFYKAVANAAACLPILLHNKKKREIQRY-EDSHQILS 500

QY 110 E-----DKSKRLNTLTNT----- 122

Db 501 QTSHDLNECSMNLITLAINKPQNKCDLDRFVNVAKTVPDDAKQLTITNTNAEALFRPG 560
 QY 123 -----MSTIY-----STGKYCNP-DNPQEC--LLPFGINEIMA---NSLDY 158
 Db 561 PGSILHLKNGPESIMNSTEYPHGSOGOLLHPGDHKAQAQHNKALPPGLSKEDAPDCSSDG 620
 QY 159 NERLWAME-----SWRSEVGRQLRPLEYEEYVVLKNEMARANHHEDYGDYWRGDEYNGV 212
 Db 621 SERSMWDDYDVYHLOGKEFEFERQCKELLEKENIMQNKQMLEHHQ----- 665
 QY 213 DGYDSRGQLIDVDEHTFEERKPLYEHLHAYVRAKLMNAYPSYISPGCLPAHLGDMKG 272
 Db 666 ----LSQFOLLQ-----EITKPVENDISKW---KPSQSLPTTNSGVSAODROLLCFYYD 713
 QY 273 RFWTNTXSLTVFPGCKPNIDVTDAMVDQAMQO--RIFKEAEKF-----FVSGGLPNM 323
 Db 714 QCFTHFTSL-----LNALDALFSCVSSAOPPRITVAHSKYVILSAHKLVIIG-DTL 763
 QY 324 TQGFWENSMLTDPGNVQAKVCHPTAWDLGKGFRIIMCTKYTMDF--LTAHHEGHIQY 381
 Db 764 TRQVTAQDITRNKVMSSNQLCEQLK-----TIVMATKMAALHPSTTALQEWVHQVT 815
 QY 382 DMAYAAQPF 390
 Db 816 DLSRNAQLF 824

Search completed: March 13, 2003, 16:58:30

Job time : 29 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 16:58:07 ; Search time 17 Seconds
(without alignments)
2182.595 Million cell updates/sec

Title: US-09-978-385-2

Perfect score: 4291

Sequence: 1 MSSSSWLLSLVAVTAQST.....ISKGNPNPQNTDDVQTSF 805

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PTCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/PTCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	4291	100.0	805	US-09-978-385-2	Sequence 2, App11
2	4287	99.9	805	US-10-114-893-86	Sequence 86, App1
3	3775	88.0	711	US-09-969-384-13	Sequence 13, App1
4	3643	84.9	681	US-09-969-384-25	Sequence 25, App1
5	3579	83.4	805	US-09-978-385-6	Sequence 6, App1
6	3561	83.0	805	US-09-978-385-9	Sequence 9, App1
7	2879	69.4	555	US-10-028-072-72	Sequence 72, App1
8	2879	69.4	555	US-10-121-049-72	Sequence 72, App1
9	2879	69.4	555	US-10-123-904-72	Sequence 72, App1
10	2879	69.4	555	US-10-140-470-72	Sequence 72, App1
11	2879	69.4	555	US-10-175-746-72	Sequence 72, App1
12	2879	69.4	555	US-10-176-918-72	Sequence 72, App1
13	2879	69.4	555	US-10-176-921-72	Sequence 72, App1
14	2879	69.4	555	US-10-137-865-72	Sequence 72, App1
15	2879	69.4	555	US-10-140-474-72	Sequence 72, App1
16	2879	69.4	555	US-10-142-431-72	Sequence 72, App1
17	2879	69.4	555	US-10-143-114-72	Sequence 72, App1
18	2879	69.4	555	US-10-140-002-72	Sequence 72, App1
19	2879	69.4	555	US-10-142-419-72	Sequence 72, App1

20	1359	31.7	261	9	US-09-969-384-23	Sequence 23, App1
21	1337	31.2	1265	9	US-09-964-899-19	Sequence 19, App1
22	990	23.1	694	10	US-09-846-996A-1	Sequence 1, App1
23	376	8.8	212	9	US-09-992-598-387	Sequence 387, App
24	376	8.8	212	9	US-09-889-292A-387	Sequence 387, App
25	376	8.8	212	9	US-09-889-735-387	Sequence 387, App
26	376	8.8	212	9	US-09-990-444-387	Sequence 387, App
27	376	8.8	212	9	US-09-989-730-387	Sequence 387, App
28	376	8.8	212	9	US-09-990-436-387	Sequence 387, App
29	376	8.8	212	9	US-09-991-181-387	Sequence 387, App
30	376	8.8	212	9	US-09-893-687-387	Sequence 387, App
31	376	8.8	212	9	US-09-889-734-387	Sequence 387, App
32	376	8.8	212	9	US-10-028-072-482	Sequence 482, App
33	376	8.8	212	9	US-09-997-653-387	Sequence 387, App
34	376	8.8	212	9	US-09-993-667-387	Sequence 387, App
35	376	8.8	212	9	US-10-121-049-482	Sequence 482, App
36	376	8.8	212	9	US-10-123-904-482	Sequence 482, App
37	376	8.8	212	9	US-10-140-470-482	Sequence 482, App
38	376	8.8	212	9	US-09-990-438-387	Sequence 387, App
39	376	8.8	212	9	US-09-990-562-387	Sequence 387, App
40	376	8.8	212	9	US-09-997-428-387	Sequence 387, App
41	376	8.8	212	9	US-09-997-666-387	Sequence 387, App
42	376	8.8	212	9	US-10-175-746-482	Sequence 482, App
43	376	8.8	212	9	US-10-176-918-482	Sequence 482, App
44	376	8.8	212	9	US-10-176-921-482	Sequence 482, App
45	376	8.8	212	9	US-09-990-711-387	Sequence 387, App

ALIGNMENTS

RESULT 1
US-09-978-385-2
Sequence 2, Application US/09978385
Patient No. US20020177211A1
GENERAL INFORMATION:
APPLICANT: Piddington, Christopher S.
APPLICANT: Petrie, Charles
APPLICANT: Shoemaker, Kimberly E.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: ZAC2: A HUMAN METALLOENZYME
FILE REFERENCE: 99-24CI
CURRENT APPLICATION NUMBER: US/09/978,385.
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 60/133,952
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: 60/151,181
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: 09/563,516
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 805
TYPE: PRT
ORGANISM: Homo sapiens
US-09-978-385-2
Query Match 100.0%; Score 4291; DB 9; Length 805;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSSSSWLLSLVAVTAQSTIEQAKFTLDFKFNHEADLFYQSSLSAAMNTITPENVO 60
DB 1 MSSSSWLLSLVAVTAQSTIEQAKFTLDFKFNHEADLFYQSSLSAAMNTITPENVO 60
QY 61 MNMNAAGKSAFLKEOSTLAQNTPLQEIQLTVKQLQALQNGSSVLSDEKSKRLNTLL 120
DB 61 MNMNAAGKSAFLKEOSTLAQNTPLQEIQLTVKQLQALQNGSSVLSDEKSKRLNTLL 120
QY 121 NTWSTVSTGKVCNPDNDQCELLLEPGLNEMNSIDYNERLWAMSWSEVGVKOLRPLY 180
DB 121 NTWSTVSTGKVCNPDNDQCELLLEPGLNEMNSIDYNERLWAMSWSEVGVKOLRPLY 180

```

QY 181 EBYVVLKEMARANHEDYGDYWRGDEYVNGVDGYDSKQQLIEDVHEHTEEEKPLYEHL 240
D 181 EBYVVLKEMARANHEDYGDYWRGDEYVNGVDGYDSKQQLIEDVHEHTEEEKPLYEHL 240
QY 241 HAYYRAKLMMNAYPSYISPIGCLPAHLGLDMMGRFWTLNLSLTVPGQKPNIDVTDAWVQ 300
D 241 HAYYRAKLMMNAYPSYISPIGCLPAHLGLDMMGRFWTLNLSLTVPGQKPNIDVTDAWVQ 300
QY 301 AMDAORIFKEAEKFFVSVGLPNNMTOGFWENSMITDPGNVQKAVCHPAMDLGKDFRILM 360
D 301 AMDAORIFKEAEKFFVSVGLPNNMTOGFWENSMITDPGNVQKAVCHPAMDLGKDFRILM 360
QY 361 CTKYTMDFLTAAHMHGHIQYDMAAQAOPFLLRNGANGEGHEAVGEIMLSAATPKHLKS 420
D 361 CTKYTMDFLTAAHMHGHIQYDMAAQAOPFLLRNGANGEGHEAVGEIMLSAATPKHLKS 420
QY 421 IGLSPDFOEDNETEINFLKQALTYGTLPFTYMLEKRWMPFKGEIPKQDMKKRWEM 480
D 421 IGLSPDFOEDNETEINFLKQALTYGTLPFTYMLEKRWMPFKGEIPKQDMKKRWEM 480
QY 481 KREIVGVPEVPYHDETCDPASLFFHVSNDYSEIRYTRTLXOFQFOEALCOAAKHGEPHL 540
D 481 KREIVGVPEVPYHDETCDPASLFFHVSNDYSEIRYTRTLXOFQFOEALCOAAKHGEPHL 540
QY 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVYGAKNMVRPLNTEPELFTWLKDQNK 600
D 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVYGAKNMVRPLNTEPELFTWLKDQNK 600
QY 601 NSFVGMSTDMSPYADOSIKVRISLSKSGKAYEMNDENKYLFRSSVAYAMROYFLKVK 660
D 601 NSFVGMSTDMSPYADOSIKVRISLSKSGKAYEMNDENKYLFRSSVAYAMROYFLKVK 660
QY 661 OMILFGEEDVRVANKPRISFNFEVYAPKAVSDIIPTEVEKAIKRSRSRINDAFRLND 720
D 661 OMILFGEEDVRVANKPRISFNFEVYAPKAVSDIIPTEVEKAIKRSRSRINDAFRLND 720
QY 721 SLEFLGIQPTLGPDPNPVSIWLVGVMGVIVGIVLIFTGIRDRKKKARSSENP 780
D 721 SLEFLGIQPTLGPDPNPVSIWLVGVMGVIVGIVLIFTGIRDRKKKARSSENP 780
QY 781 YASIDISKGENNPGFQNTDVQTSF 805
D 781 YASIDISKGENNPGFQNTDVQTSF 805

RESULT 2
US-10-114-893-86
; Sequence 86, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalite, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-104
; CURRENT APPLICATION NUMBER: US/10/114,893
; EARLIER FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 86
; LENGTH: 805

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-893-86
Query Match 99.9%; Score 4287; DB 9; Length 805;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 804; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESSWLLSLVAVTAAGSTIEQAKTFPLDKRNHAEDELFOSSLASNNYNTITEENVQ 60
D 1 MESSWLLSLVAVTAAGSTIEQAKTFPLDKRNHAEDELFOSSLASNNYNTITEENVQ 60
QY 61 MNMNGDKWSAEFLKQSTLAQWYPIQETQNLTVKLOLQALQONSSVSEDSKRLNTIL 120
D 61 MNMNGDKWSAEFLKQSTLAQWYPIQETQNLTVKLOLQALQONSSVSEDSKRLNTIL 120
QY 121 NTMSTIYSTGKVCNPDNPQECILLEPGINEIMANSLDYNERLMAESWRSEYQOLRPLY 180
D 121 NTMSTIYSTGKVCNPDNPQECILLEPGINEIMANSLDYNERLMAESWRSEYQOLRPLY 180
QY 181 EBYVVLKEMARANHEDYGDYWRGDEYVNGVDGYDSKQQLIEDVHEHTEEEKPLYEHL 240
D 181 EBYVVLKEMARANHEDYGDYWRGDEYVNGVDGYDSKQQLIEDVHEHTEEEKPLYEHL 240
QY 241 HAYYRAKLMMNAYPSYISPIGCLPAHLGLDMMGRFWTLNLSLTVPGQKPNIDVTDAWVQ 300
D 241 HAYYRAKLMMNAYPSYISPIGCLPAHLGLDMMGRFWTLNLSLTVPGQKPNIDVTDAWVQ 300
QY 301 AMDAORIFKEAEKFFVSVGLPNNMTOGFWENSMITDPGNVQKAVCHPAMDLGKDFRILM 360
D 301 AMDAORIFKEAEKFFVSVGLPNNMTOGFWENSMITDPGNVQKAVCHPAMDLGKDFRILM 360
QY 361 CTKYTMDFLTAAHMHGHIQYDMAAQAOPFLLRNGANGEGHEAVGEIMLSAATPKHLKS 420
D 361 CTKYTMDFLTAAHMHGHIQYDMAAQAOPFLLRNGANGEGHEAVGEIMLSAATPKHLKS 420
QY 421 IGLSPDFOEDNETEINFLKQALTYGTLPFTYMLEKRWMPFKGEIPKQDMKKRWEM 480
D 421 IGLSPDFOEDNETEINFLKQALTYGTLPFTYMLEKRWMPFKGEIPKQDMKKRWEM 480
QY 481 KREIVGVPEVPYHDETCDPASLFFHVSNDYSEIRYTRTLXOFQFOEALCOAAKHGEPHL 540
D 481 KREIVGVPEVPYHDETCDPASLFFHVSNDYSEIRYTRTLXOFQFOEALCOAAKHGEPHL 540
QY 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVYGAKNMVRPLNTEPELFTWLKDQNK 600
D 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVYGAKNMVRPLNTEPELFTWLKDQNK 600
QY 601 NSFVGMSTDMSPYADOSIKVRISLSKSGKAYEMNDENKYLFRSSVAYAMROYFLKVK 660
D 601 NSFVGMSTDMSPYADOSIKVRISLSKSGKAYEMNDENKYLFRSSVAYAMROYFLKVK 660
QY 661 OMILFGEEDVRVANKPRISFNFEVYAPKAVSDIIPTEVEKAIKRSRSRINDAFRLND 720
D 661 OMILFGEEDVRVANKPRISFNFEVYAPKAVSDIIPTEVEKAIKRSRSRINDAFRLND 720
QY 721 SLEFLGIQPTLGPDPNPVSIWLVGVMGVIVGIVLIFTGIRDRKKKARSSENP 780
D 721 SLEFLGIQPTLGPDPNPVSIWLVGVMGVIVGIVLIFTGIRDRKKKARSSENP 780
QY 781 YASIDISKGENNPGFQNTDVQTSF 805
D 781 YASIDISKGENNPGFQNTDVQTSF 805

RESULT 3
US-09-969-384-13
; Sequence 13, Application US/09969384
; Publication No. US20020192749A1
; GENERAL INFORMATION:
; APPLICANT: Moore, et al.
; TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT055P1

```

CURRENT APPLICATION NUMBER: US/09/969,384
CURRENT FILING DATE: 2001-10-03
PRIOR APPLICATION NUMBER: PCT/US01/10542
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 60/236,384
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/194,118
PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 711
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (219)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (240)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (499)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-969-384-13

Query Match 88.0%; Score 3775; DB 9; Length 711;
Best Local Similarity 99.3%; Pred. No. 1.8e-305;
Matches 704; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 62 MNNAGDKWSAFLEKQSTLAQMYPLQEIQNTLVKQLQALQONGSSVLESDSKRLNTILN 121
DB 1 MNNAGDKWSAFLEKQSTLAQMYPLQEIQNTLVKQLQALQONGSSVLESDSKRLNTILN 60
QY 122 TMSITSTGKVCNPDNPOECILLEPGLNEIMANSIDYNERLMAWESRSEVQKQRLPYE 181
DB 61 TMSITSTGKVCNPDNPOECILLEPGLNEIMANSIDYNERLMAWESRSEVQKQRLPYE 120
QY 182 EYVVLAKNEARAHNEYEDGYWRGDEYNGVDYDSRGQLEDEYHTFEELIKPLYEHL 241
DB 121 EYVVLAKNEARAHNEYEDGYWRGDEYNGVDYDSRGQLEDEYHTFEELIKPLYEHL 180
QY 242 AYVRKIMNAVPSYISPIGCLPAHLGDMGRFWNTNLSLTVPGQKRNIDVTDAVDQA 301
DB 181 AYVRKIMNAVPSYISPIGCLPAHLGDMGRFWNTNLSLTVPGQKRNIDVTDAVDQA 240
QY 302 WDAQRIFEAEKEFVSVGLPNTQGFWENSMILDPGNVQKAVCHPTAMDLSKGFRLIMC 361
DB 241 WDAQRIFEAEKEFVSVGLPNTQGFWENSMILDPGNVQKAVCHPTAMDLSKGFRLIMC 300
QY 362 TKTVMDDFLTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMLSAATPKHLKSI 421
DB 301 TKTVMDDFLTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMLSAATPKHLKSI 360
QY 422 GLLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWAVFEGEIPKQMMKKWEMK 481
DB 361 GLLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWAVFEGEIPKQMMKKWEMK 420
QY 482 RELVGVVEPVPHDETCYDPAFLFHVSNDSYFIRYTRTYLOFOFQALQOAKHNEGRLHK 541
DB 421 RELVGVVEPVPHDETCYDPAFLFHVSNDSYFIRYTRTYLOFOFQALQOAKHNEGRLHK 480
QY 542 CDSINSTEAGOKLFNMLRLCKSEPTLALENVGAKNMVRLPLNFEELFTWLKQONKN 601
DB 481 CDSINSTEAGOKLFNMLRLCKSEPTLALENVGAKNMVRLPLNFEELFTWLKQONKN 540
QY 602 SPVGMSTWSPYADQSIKVRISLKSALGDKAYEMDNENKYLFRSSVAYVAMROYETLKYKQ 661
DB 541 SPVGMSTWSPYADQSIKVRISLKSALGDKAYEMDNENKYLFRSSVAYVAMROYETLKYKQ 600
QY 662 MLLFGEDVAVANLKPRISFNFTVTPAKNSDIIIPTEVEKATIRMSRSRIINDAFRLDNS 721
DB 601 MLLFGEDVAVANLKPRISFNFTVTPAKNSDIIIPTEVEKATIRMSRSRIINDAFRLDNS 660

QY 722 LFLGIQPTLGPNNQPPVSIWLIVFGVWGVIVGIVILIFTGRDRK 770
DB 661 LFLGIQPTLGPNNQPPVSIWLIVFGVWGVIVGIVILIFTGRDRK 709

RESULT 4
US-09-969-384-25
Sequence 25, Application US/09969384
Publication No. US20020192749A1
GENERAL INFORMATION:
APPLICANT: Moore, et al.
TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
FILE REFERENCE: PT055P1
CURRENT APPLICATION NUMBER: US/09/969,384
CURRENT FILING DATE: 2001-10-03
PRIOR APPLICATION NUMBER: PCT/US01/10542
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 60/236,384
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/194,118
PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 25
LENGTH: 681
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (219)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (240)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (499)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-969-384-25

Query Match 84.9%; Score 3643; DB 9; Length 681;
Best Local Similarity 99.3%; Pred. No. 1.6e-294;
Matches 676; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 62 MNNAGDKWSAFLEKQSTLAQMYPLQEIQNTLVKQLQALQONGSSVLESDSKRLNTILN 121
DB 1 MNNAGDKWSAFLEKQSTLAQMYPLQEIQNTLVKQLQALQONGSSVLESDSKRLNTILN 60
QY 122 TMSITSTGKVCNPDNPOECILLEPGLNEIMANSIDYNERLMAWESRSEVQKQRLPYE 181
DB 61 TMSITSTGKVCNPDNPOECILLEPGLNEIMANSIDYNERLMAWESRSEVQKQRLPYE 120
QY 182 EYVVLAKNEARAHNEYEDGYWRGDEYNGVDYDSRGQLEDEYHTFEELIKPLYEHL 241
DB 121 EYVVLAKNEARAHNEYEDGYWRGDEYNGVDYDSRGQLEDEYHTFEELIKPLYEHL 180
QY 242 AYVRKIMNAVPSYISPIGCLPAHLGDMGRFWNTNLSLTVPGQKRNIDVTDAVDQA 301
DB 181 AYVRKIMNAVPSYISPIGCLPAHLGDMGRFWNTNLSLTVPGQKRNIDVTDAVDQA 240
QY 302 WDAQRIFEAEKEFVSVGLPNTQGFWENSMILDPGNVQKAVCHPTAMDLSKGFRLIMC 361
DB 241 WDAQRIFEAEKEFVSVGLPNTQGFWENSMILDPGNVQKAVCHPTAMDLSKGFRLIMC 300
QY 362 TKTVMDDFLTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMLSAATPKHLKSI 421
DB 301 TKTVMDDFLTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMLSAATPKHLKSI 360
QY 422 GLLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWAVFEGEIPKQMMKKWEMK 481
DB 361 GLLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWAVFEGEIPKQMMKKWEMK 420
QY 482 RELVGVVEPVPHDETCYDPAFLFHVSNDSYFIRYTRTYLOFOFQALQOAKHNEGRLHK 541

```

Db 421 REIVGVEPEVPHDETCDPASLPHVSNDSFIYRTLYQFOFALCOQAKHCEPLK 480
QY 542 CDISNSTEAGOKLFNNLRGKSEPMWLALENVGAKNMVRPLNFEPLFTWLKDQKN 601
Db 481 CDISNSTEAGOKLFNNLRGKSEPMWLALENVGAKNMVRPLNFEPLFTWLKDQKN 540
QY 602 SEYGMSTWSPYADOSIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVKNO 661
Db 541 SEYGMSTWSPYADOSIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVKNO 600
QY 662 MLFGEDEVANLKRISFNFEVYAPKAVSDIIPREVEKAIKMSRSRINDAFRLDNS 721
Db 601 MLFGEDEVANLKRISFNFEVYAPKAVSDIIPREVEKAIKMSRSRINDAFRLDNS 660
QY 722 LEFLGIQPTLGPPOPPVSIW 742
Db 661 LEFLGIQPTLGPPOPPVSIW 681

```

RESULT 5

```

US-09-978-385-6
; Sequence 6, Application US/09978385
; Patent No. US20020177211A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Pectie, Charles
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME
; FILE REFERENCE: 99-24C1
; CURRENT APPLICATION NUMBER: US/09/978,385
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/133,952
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 60/151,181
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/563,516
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 6
; LENGTH: 805
; TYPE: PRP
; ORGANISM: Mouse
US-09-978-385-6

```

```

Query Match 83.4%; Score 3579; DB 9; Length 805;
Best Local Similarity 82.1%; Pred. No. 4.4e-289;
Matches 661; Conservative 60; Mismatches 84; Indels 0; Gaps 0;

```

```

QY 1 MSSSSWLLSLVAVTAOSTIEQAKTFLEKFNHEADLEFYOSSLASWNTNITEENVQ 60
Db 1 MSSSSWLLSLVAVTAOSTIEQAKTFLEKFNHEADLEFYOSSLASWNTNITEENVQ 60
QY 61 MNNAAGDKWSAFLEKQOSTLAQMYPLQEIQNTLVKLOALQOQSSVLSDDKSKRLNTIL 120
Db 61 KMSEAAKWSAFYEQSKTQOSFSLQEIQPIIKROLQALQOQSSVLSDDKSKRLNTIL 120
QY 121 NTWSTIYSTGVCPNDPQECLELLEPGLNEIMANSIDYNERLWAMSRSEVQKOLRPY 180
Db 121 NTWSTIYSTGVCPNDPQECLELLEPGLNEIMANSIDYNERLWAMSRSEVQKOLRPY 180
QY 181 EEVYVLKNEBARANHYEDGDYRGDYEVANGVGYISGQGLIEDVEHTFEETKPLYEHL 240
Db 181 EEVYVLKNEBARANHYEDGDYRGDYEVANGVGYISGQGLIEDVEHTFEETKPLYEHL 240
QY 241 HAVYRAKLNMAVPSYISPLGCLPAHLGLDMGGRFTNLSLVPEQOKNIDVTAMVQ 300
Db 241 HAVYRRKIMDTYSYISPLGCLPAHLGLDMGGRFTNLSLVPEQOKNIDVTAMVQ 300
QY 301 AMDAOKIFEAKEFFVSGPLNMTQGFWNSMLTDPGNVQKAVCHPTAMDLCGDFRILM 360
Db 301 AMDAOKIFEAKEFFVSGPLNMTQGFWNSMLTDPGNVQKAVCHPTAMDLCGDFRILM 360

```

```

Db 301 GMDAERIQEAKKEFFVSGPLNMTQGFWNSMLTDPADGRKVCCHPTAMDLCGDFRILM 360
QY 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAOPELLRNGANGEGFEAVGELMISLSAATPKILKS 420
Db 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAOPELLRNGANGEGFEAVGELMISLSAATPKILKS 420
QY 421 IGLSPQOEENETEFNLKQALITVGLPEFTYMLKRWMYFKGIPRDKMKWEM 480
Db 421 IGLSPQOEENETEFNLKQALITVGLPEFTYMLKRWMYFKGIPRDKMKWEM 480
QY 481 KREIVGVEPEVPHDETCDPASLPHVSNDSFIYRTLYQFOFALCOQAKHCEPLK 540
Db 481 KREIVGVEPEVPHDETCDPASLPHVSNDSFIYRTLYQFOFALCOQAKHCEPLK 540
QY 541 KCDISNSTEAGOKLFNNLRGKSEPMWLALENVGAKNMVRPLNFEPLFTWLKDQKN 600
Db 541 KCDISNSTEAGOKLFNNLRGKSEPMWLALENVGAKNMVRPLNFEPLFTWLKDQKN 600
QY 601 NSFVGMSTWSPYADOSIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVKNO 660
Db 601 NSFVGMSTWSPYADOSIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVKNO 660
QY 661 QMLFGEDEVANLKRISFNFEVYAPKAVSDIIPREVEKAIKMSRSRINDAFRLDNS 720
Db 661 QVPELEEDYRVGDLKRVSYFFVTSPOVSDVIFRSEVEDAIRMSRGRINDVGLDND 720
QY 721 SLEFLGIQPTLGPPOPPVSIWLVFVGWGVIVGLVILFTGIDRRKKKKARSGENP 780
Db 721 SLEFLGIHPLLEPPYQPPVITWILIFGVMAVAVGIIILVYTGIRKKKKETKRENP 780
QY 781 YASIDISGENNPGFQNTDVOYTSF 805
Db 781 YDSMDIGKGSNAGFQNSDDAQTSTF 805

```

RESULT 6

```

US-09-978-385-9
; Sequence 9, Application US/09978385
; Patent No. US20020177211A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Pectie, Charles
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME
; FILE REFERENCE: 99-24C1
; CURRENT APPLICATION NUMBER: US/09/978,385
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/133,952
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 60/151,181
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/563,516
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 9
; LENGTH: 805
; TYPE: PRP
; ORGANISM: Mouse
US-09-978-385-9

```

```

Query Match 83.0%; Score 3561; DB 9; Length 805;
Best Local Similarity 81.9%; Pred. No. 1.4e-287;
Matches 659; Conservative 60; Mismatches 86; Indels 0; Gaps 0;

```

```

QY 1 MSSSSWLLSLVAVTAOSTIEQAKTFLEKFNHEADLEFYOSSLASWNTNITEENVQ 60
Db 1 MSSSSWLLSLVAVTAOSTIEQAKTFLEKFNHEADLEFYOSSLASWNTNITEENVQ 60
QY 61 MNNAAGDKWSAFLEKQOSTLAQMYPLQEIQNTLVKLOALQOQSSVLSDDKSKRLNTIL 120
Db 61 KMSEAAKWSAFYEQSKTQOSFSLQEIQPIIKROLQALQOQSSVLSDDKSKRLNTIL 120

```

QY 121 NTMSTIYSGKVCNPDNPOECILLEPGLNEMANSLDYNERLWAMESWSENGKOLRPLY 180
121 NTMSTIYSGKVCNPDNPOECILLEPGLNEMANSLDYNERLWAMESWSENGKOLRPLY 180
Db 121 NTMSTIYSGKVCNPDNPOECILLEPGLNEMANSLDYNERLWAMESWSENGKOLRPLY 180
QY 181 EEEVVLKNEBARNHEDYDYGRCYGVNGVGYRSROLLLEDVHPEEIKPLYEH 240
181 EEEVVLKNEBARNHEDYDYGRCYGVNGVGYRSROLLLEDVHPEEIKPLYEH 240
Db 181 EEEVVLKNEBARNHEDYDYGRCYGVNGVGYRSROLLLEDVHPEEIKPLYEH 240
QY 241 HAYRAKIMANAPSYISPIGLPAHLILGDMWGRFTNLXSLTVPFQKPNIDVTDAVDO 300
241 HAYRAKIMANAPSYISPIGLPAHLILGDMWGRFTNLXSLTVPFQKPNIDVTDAVDO 300
Db 241 HAYRAKIMANAPSYISPIGLPAHLILGDMWGRFTNLXSLTVPFQKPNIDVTDAVDO 300
QY 301 AMDAQRIFKEAEKFFVYVGLPNNYQGWEMSLTDGNNVQKACHPAMWLGKDRPILM 360
301 AMDAQRIFKEAEKFFVYVGLPNNYQGWEMSLTDGNNVQKACHPAMWLGKDRPILM 360
Db 301 AMDAQRIFKEAEKFFVYVGLPNNYQGWEMSLTDGNNVQKACHPAMWLGKDRPILM 360
QY 361 CTKTMDPFLAHHEMHIQYDAAAYAPFLIRNGANECHHEAVGELMSLSAATPKHLKS 420
361 CTKTMDPFLAHHEMHIQYDAAAYAPFLIRNGANECHHEAVGELMSLSAATPKHLKS 420
Db 361 CTKTMDPFLAHHEMHIQYDAAAYAPFLIRNGANECHHEAVGELMSLSAATPKHLKS 420
QY 421 IGLSPDEQEDNETEINFLLKQALTYGLPFTYMLEKRWVFKGEIPKDOMKKWEM 480
421 IGLSPDEQEDNETEINFLLKQALTYGLPFTYMLEKRWVFKGEIPKDOMKKWEM 480
Db 421 IGLSPDEQEDNETEINFLLKQALTYGLPFTYMLEKRWVFKGEIPKDOMKKWEM 480
QY 481 KREIVGYEPYHDETCDPASLEFHVNDYSEFITYRTLYXOPQOFALQAAKHGEPH 540
481 KREIVGYEPYHDETCDPASLEFHVNDYSEFITYRTLYXOPQOFALQAAKHGEPH 540
Db 481 KREIVGYEPYHDETCDPASLEFHVNDYSEFITYRTLYXOPQOFALQAAKHGEPH 540
QY 541 KCDISNSTEAGOKLFNNMLRGLKSEPTLAEENVGAKNMWRPLNFEPEFTLWLNQNK 600
541 KCDISNSTEAGOKLFNNMLRGLKSEPTLAEENVGAKNMWRPLNFEPEFTLWLNQNK 600
Db 541 KCDISNSTEAGOKLFNNMLRGLKSEPTLAEENVGAKNMWRPLNFEPEFTLWLNQNK 600
QY 601 NSFVGMWTEWSPYADQSIKVRISLKSALGDAYENDNEMTLFRSSVAYAMROYFLKYKN 660
601 NSFVGMWTEWSPYADQSIKVRISLKSALGDAYENDNEMTLFRSSVAYAMROYFLKYKN 660
Db 601 NSFVGMWTEWSPYADQSIKVRISLKSALGDAYENDNEMTLFRSSVAYAMROYFLKYKN 660
QY 661 QMIFGEDVYANIKPRISNFEVTPAKNVDIIPREVEKAIIRMSRSRINDPFLNDN 720
661 QMIFGEDVYANIKPRISNFEVTPAKNVDIIPREVEKAIIRMSRSRINDPFLNDN 720
Db 661 QMIFGEDVYANIKPRISNFEVTPAKNVDIIPREVEKAIIRMSRSRINDPFLNDN 720
QY 721 SLEFIOPTLGPNOQPVSIWLVGVGMVIVGIVLITGIRPKKKKKARSGENP 780
721 SLEFIOPTLGPNOQPVSIWLVGVGMVIVGIVLITGIRPKKKKKARSGENP 780
Db 721 SLEFIOPTLGPNOQPVSIWLVGVGMVIVGIVLITGIRPKKKKKARSGENP 780
QY 781 YASIDISKGENNPFQNTDVQTSF 805
781 YASIDISKGENNPFQNTDVQTSF 805
Db 781 YASIDISKGENNPFQNTDVQTSF 805

RESULT 7
US-10-028-072-72
; Sequence 72, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tunes, Daniel
; APPLICANT: Matanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:

FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028, 072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062827
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/066770
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/069212
 PRIOR FILING DATE: 1997-12-11
 PRIOR APPLICATION NUMBER: 60/069278
 PRIOR FILING DATE: 1997-12-11
 PRIOR APPLICATION NUMBER: 60/069334
 PRIOR FILING DATE: 1997-12-11
 PRIOR APPLICATION NUMBER: 60/069694
 PRIOR FILING DATE: 1997-12-16
 PRIOR APPLICATION NUMBER: 60/072320
 PRIOR FILING DATE: 1998-01-23
 PRIOR APPLICATION NUMBER: 60/073612
 PRIOR FILING DATE: 1998-02-04
 PRIOR APPLICATION NUMBER: 60/074086
 PRIOR FILING DATE: 1998-02-09
 PRIOR APPLICATION NUMBER: 60/074092
 PRIOR FILING DATE: 1998-02-09
 PRIOR APPLICATION NUMBER: 60/077791
 PRIOR FILING DATE: 1998-03-12
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/079294
 PRIOR FILING DATE: 1998-03-25
 PRIOR APPLICATION NUMBER: 60/079663
 PRIOR FILING DATE: 1998-02-27
 PRIOR APPLICATION NUMBER: 60/079728
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/080165
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/081203
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081229
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081695
 PRIOR FILING DATE: 1998-04-14
 PRIOR APPLICATION NUMBER: 60/081817
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081818
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/082999
 PRIOR FILING DATE: 1998-04-24
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/083545
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084627
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084637
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085149
 PRIOR FILING DATE: 1998-05-12
 PRIOR APPLICATION NUMBER: 60/085323
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085338
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085339
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085697
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085704
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/086414
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/086430
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/087106
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/088026

PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088730
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088741
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088810
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090538
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090863
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/091360
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091519
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091982
 PRIOR FILING DATE: 1998-07-07

Query Match 69.4%; Score 2979; DB 9; Length 555;
 Best Local Similarity 99.8%; Pred. No. 1.9e-239;
 Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M S S S W L L S I V A V T A O S T I E O A K T F L D K E N H A E D L F Y O S S I A S N Y N T N T E E N V Q 60
 Db 1 M S S S W L L S I V A V T A O S T I E O A K T F L D K E N H A E D L F Y O S S I A S N Y N T N T E E N V Q 60
 QY 61 N N N A G D K W S A F L K E O S T L A O M Y P L O E I O N L T V K L Q L A L O Q N S S V S E D K S K R L N T I L 120
 Db 61 N N N A G D K W S A F L K E O S T L A O M Y P L O E I O N L T V K L Q L A L O Q N S S V S E D K S K R L N T I L 120
 QY 121 N M S T I Y S T G K V C N P D N Q E C L L E P G L N E T A N S L D Y N E R I A M E S R S E V G K O L R P L Y 180
 Db 121 N M S T I Y S T G K V C N P D N Q E C L L E P G L N E T A N S L D Y N E R I A M E S R S E V G K O L R P L Y 180
 QY 181 E E Y V V L K N E M A R A N H Y E D G D Y W R G D Y E V N G V D Y S R G O L I E D V E H T F E E I P L Y E H L 240
 Db 181 E E Y V V L K N E M A R A N H Y E D G D Y W R G D Y E V N G V D Y S R G O L I E D V E H T F E E I P L Y E H L 240
 QY 241 H A Y V R A K L M A N A P S Y I S P I G C L P A H L L G D M G R E W T N L Y S L T V P G O K P N I D V T D A M V Q 300
 Db 241 H A Y V R A K L M A N A P S Y I S P I G C L P A H L L G D M G R E W T N L Y S L T V P G O K P N I D V T D A M V Q 300
 QY 301 A M D A O R I F K E A K E F P V S G L P M T O G F E W E N S M L D P G V O A V C H P T A M D I G K D F R I L M 360
 Db 301 A M D A O R I F K E A K E F P V S G L P M T O G F E W E N S M L D P G V O A V C H P T A M D I G K D F R I L M 360
 QY 361 C K V T M D F L T A H H E M G I O Y D M A Y A O P F L R N G A N G F E A V G E I M S I S A A T P K H L K S 420
 Db 361 C K V T M D F L T A H H E M G I O Y D M A Y A O P F L R N G A N G F E A V G E I M S I S A A T P K H L K S 420
 QY 421 I G L S P D F O E D N E T E I N F L K O A L T I V G T L P F T Y M L E K W M V E K G E I P K D W M K K W E M 480
 Db 421 I G L S P D F O E D N E T E I N F L K O A L T I V G T L P F T Y M L E K W M V E K G E I P K D W M K K W E M 480
 QY 481 K R E I G V V E P V P H D E T Y C D P A S L F H V S D Y S F I R Y T R T L Y O F O P O A L C O A K H E G P L H 540
 Db 481 K R E I G V V E P V P H D E T Y C D P A S L F H V S D Y S F I R Y T R T L Y O F O P O A L C O A K H E G P L H 540

OY 541 KCDISNSTEAGOKL 554
DB 541 KCDISNSTEAGOKL 554

RESULT 8

US-10-121-049-72
Sequence 72, Application US/10121049
Publication No. US20030022239A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Mei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C17
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-049-72

Query Match 69.4%; Score 2979; DB 5; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSSSSWLLSLVAVTAOSTIEEOAKTFLDKFNHEAEDLFYQSSSLASWNTNITEENVQ 60
DB 1 MSSSSWLLSLVAVTAOSTIEEOAKTFLDKFNHEAEDLFYQSSSLASWNTNITEENVQ 60
OY 61 NNNNAGDKWSAFLEKQSTLAQMPLOEIONLVKIQLOALQONGSSVLESDSKRLNTTL 120
DB 61 NNNNAGDKWSAFLEKQSTLAQMPLOEIONLVKIQLOALQONGSSVLESDSKRLNTTL 120
OY 121 NTMSTIYTGKVCNPNPOECILLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPLY 180
DB 121 NTMSTIYTGKVCNPNPOECILLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPLY 180
OY 181 EEEVVLKNEMARANHYEDYGYWRGDEYVNGVDYDYSRGOLIEDVEHTEFEIKPLYEHL 240
DB 181 EEEVVLKNEMARANHYEDYGYWRGDEYVNGVDYDYSRGOLIEDVEHTEFEIKPLYEHL 240
OY 241 HAYYRAKLMNAYPSYISPIGCLPAHLGDMGRFTNLSLTVPGOKPNIDVTDAVQ 300
DB 241 HAYYRAKLMNAYPSYISPIGCLPAHLGDMGRFTNLSLTVPGOKPNIDVTDAVQ 300
OY 301 AMDAORIFKEAEKFFVSGLPNMTQGFWENSMLTDPGNVOKAVCHPTAMDIGKGFRIIM 360
DB 301 AMDAORIFKEAEKFFVSGLPNMTQGFWENSMLTDPGNVOKAVCHPTAMDIGKGFRIIM 360
OY 361 CTKYTMDDFLTAHHEMGIQYDAAYAAQPFLLRNGANEHGEHVAAGETIMSLSAATPKHLKS 420
DB 361 CTKYTMDDFLTAHHEMGIQYDAAYAAQPFLLRNGANEHGEHVAAGETIMSLSAATPKHLKS 420

OY 421 IGLSPDFQEDNETEINFLKQALTIYGTLPPTYMLEKRWNVKGEIPKDKWKKWEM 480
DB 421 IGLSPDFQEDNETEINFLKQALTIYGTLPPTYMLEKRWNVKGEIPKDKWKKWEM 480
OY 481 KREIVGVEPVPHDETCDDPASLFHVSNDYSFIRYRTYTLQFOFQALCOAAKHGEPFLH 540
DB 481 KREIVGVEPVPHDETCDDPASLFHVSNDYSFIRYRTYTLQFOFQALCOAAKHGEPFLH 540
OY 541 KCDISNSTEAGOKL 554
DB 541 KCDISNSTEAGOKL 554

RESULT 9

US-10-123-904-72
Sequence 72, Application US/10123904
Publication No. US20030022328A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Mei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C54
CURRENT FILING DATE: 2002-04-16
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-123-904-72

Query Match 69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSSSSWLLSLVAVTAOSTIEEOAKTFLDKFNHEAEDLFYQSSSLASWNTNITEENVQ 60
DB 1 MSSSSWLLSLVAVTAOSTIEEOAKTFLDKFNHEAEDLFYQSSSLASWNTNITEENVQ 60
OY 61 NNNNAGDKWSAFLEKQSTLAQMPLOEIONLVKIQLOALQONGSSVLESDSKRLNTTL 120
DB 61 NNNNAGDKWSAFLEKQSTLAQMPLOEIONLVKIQLOALQONGSSVLESDSKRLNTTL 120
OY 121 NTMSTIYTGKVCNPNPOECILLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPLY 180
DB 121 NTMSTIYTGKVCNPNPOECILLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPLY 180
OY 181 EEEVVLKNEMARANHYEDYGYWRGDEYVNGVDYDYSRGOLIEDVEHTEFEIKPLYEHL 240
DB 181 EEEVVLKNEMARANHYEDYGYWRGDEYVNGVDYDYSRGOLIEDVEHTEFEIKPLYEHL 240
OY 241 HAYYRAKLMNAYPSYISPIGCLPAHLGDMGRFTNLSLTVPGOKPNIDVTDAVQ 300
DB 241 HAYYRAKLMNAYPSYISPIGCLPAHLGDMGRFTNLSLTVPGOKPNIDVTDAVQ 300
OY 301 AMDAORIFKEAEKFFVSGLPNMTQGFWENSMLTDPGNVOKAVCHPTAMDIGKGFRIIM 360

|||||
Db 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWEKSMITDPCGNVQKAVCHPTAMDIGKGRDFRILM 360
Qy 361 CTKYVMDDELTAHHEMGHIQYDMAVYAAOPFLLRNGANEGFHEAVGELMSAATPKHLKS 420
Db 361 CTKYVMDDELTAHHEMGHIQYDMAVYAAOPFLLRNGANEGFHEAVGELMSAATPKHLKS 420
Qy 421 IGLSPDPOEDNETEINFLKQALITVGTLPFTYMLEKRWMMVFKEGELPKDQMKKWMEM 480
Db 421 IGLSPDPOEDNETEINFLKQALITVGTLPFTYMLEKRWMMVFKEGELPKDQMKKWMEM 480
Qy 481 KREIVGVEPVPHEDETCDPASLFHVSNDYSFIRYTRITLYQFOFQALCOAAKHGEPHL 540
Db 481 KREIVGVEPVPHEDETCDPASLFHVSNDYSFIRYTRITLYQFOFQALCOAAKHGEPHL 540
Qy 541 KCDISNSTEAGOKL 554
Db 541 KCDISNSTEAGOKL 554

RESULT 10

US-10-140-470-72
; Sequence 72, Application US/10140470
; Publication No. US2003002231A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C160
CURRENT APPLICATION NUMBER: US/10/140,470
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-470-72

Query Match

69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;

Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSSSSWLLSLVAVTAASSTIEBOAKTFLDKFNEHAEDELFYOSSLASMNNTNTEENVQ 60
Db 1 MSSSSWLLSLVAVTAASSTIEBOAKTFLDKFNEHAEDELFYOSSLASMNNTNTEENVQ 60
Qy 61 MNMNAAGKWSAFLKEOSTLAQMPLOEIONLTVKLOLQALQONSSSVLSSEDKSKRLNTIL 120
Db 61 MNMNAAGKWSAFLKEOSTLAQMPLOEIONLTVKLOLQALQONSSSVLSSEDKSKRLNTIL 120
Qy 121 NTMSTIYSTGKVCNPNPOCCLLEPGLNEIMANSIDYNERLWAMESWRSSEVGKOLRPLY 180
Db 121 NTMSTIYSTGKVCNPNPOCCLLEPGLNEIMANSIDYNERLWAMESWRSSEVGKOLRPLY 180
Qy 181 EEFVVLAKENARANHEDYGDYWRGDEYVNGVGYDYISGQLIEDYEHFEELKPLYEHL 240
Db 181 EEFVVLAKENARANHEDYGDYWRGDEYVNGVGYDYISGQLIEDYEHFEELKPLYEHL 240

Db 181 EEFVVLAKENARANHEDYGDYWRGDEYVNGVGYDYISGQLIEDYEHFEELKPLYEHL 240
Qy 241 HAYVRAKIMNAYSYSPICGLPAHLIGMWGFNTNLSLVYFPEGQKNIDVTAMDQ 300
Db 241 HAYVRAKIMNAYSYSPICGLPAHLIGMWGFNTNLSLVYFPEGQKNIDVTAMDQ 300
Qy 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWEKSMITDPCGNVQKAVCHPTAMDIGKGRDFRILM 360
Db 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWEKSMITDPCGNVQKAVCHPTAMDIGKGRDFRILM 360
Qy 361 CTKYVMDDELTAHHEMGHIQYDMAVYAAOPFLLRNGANEGFHEAVGELMSAATPKHLKS 420
Db 361 CTKYVMDDELTAHHEMGHIQYDMAVYAAOPFLLRNGANEGFHEAVGELMSAATPKHLKS 420
Qy 421 IGLSPDPOEDNETEINFLKQALITVGTLPFTYMLEKRWMMVFKEGELPKDQMKKWMEM 480
Db 421 IGLSPDPOEDNETEINFLKQALITVGTLPFTYMLEKRWMMVFKEGELPKDQMKKWMEM 480
Qy 481 KREIVGVEPVPHEDETCDPASLFHVSNDYSFIRYTRITLYQFOFQALCOAAKHGEPHL 540
Db 481 KREIVGVEPVPHEDETCDPASLFHVSNDYSFIRYTRITLYQFOFQALCOAAKHGEPHL 540
Qy 541 KCDISNSTEAGOKL 554
Db 541 KCDISNSTEAGOKL 554

RESULT 11

US-10-175-746-72
; Sequence 72, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C353
CURRENT APPLICATION NUMBER: US/10/175,746
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-746-72

Query Match

69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;

Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSSSSWLLSLVAVTAASSTIEBOAKTFLDKFNEHAEDELFYOSSLASMNNTNTEENVQ 60
Db 1 MSSSSWLLSLVAVTAASSTIEBOAKTFLDKFNEHAEDELFYOSSLASMNNTNTEENVQ 60
Qy 61 MNMNAAGKWSAFLKEOSTLAQMPLOEIONLTVKLOLQALQONSSSVLSSEDKSKRLNTIL 120
Db 61 MNMNAAGKWSAFLKEOSTLAQMPLOEIONLTVKLOLQALQONSSSVLSSEDKSKRLNTIL 120

```

QY 121 NTMSTIYTGKVCNPDNPOECLELLEPGLINEIMANSLDYNERNLWAMESRSEVGKOLRPLY 180
    |||||||
Db 121 NMSTIYTGKVCNPDNPOECLELLEPGLINEIMANSLDYNERNLWAMESRSEVGKOLRPLY 180
QY 181 EBYVVLKNEMARANHYEDYGDYWRGDEYVNGVDYDYSRGQLIEDVEHTFEELKPLYEHL 240
    |||||||
Db 181 EBYVVLKNEMARANHYEDYGDYWRGDEYVNGVDYDYSRGQLIEDVEHTFEELKPLYEHL 240
QY 241 HAYVRAKIMANAPSYISIPICLPAHLIGDMWGRFTNLXSLTVPFGOKPNIDVTAMDYQ 300
    |||||||
Db 241 HAYVRAKIMANAPSYISIPICLPAHLIGDMWGRFTNLXSLTVPFGOKPNIDVTAMDYQ 300
QY 301 AMDAORIFKEAEKFEFVSGLPNMTQGFWENSMLTDPGNVQKAVCHPTAMDLGKGFRIIM 360
    |||||||
Db 301 AMDAORIFKEAEKFEFVSGLPNMTQGFWENSMLTDPGNVQKAVCHPTAMDLGKGFRIIM 360
QY 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAPPELLRNGANEGHEAVGEIMSLSAATPKHLKS 420
    |||||||
Db 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAPPELLRNGANEGHEAVGEIMSLSAATPKHLKS 420
QY 421 IGLSPDQEDNTEINFLKQALITVGLTPTMYLKKRMWVFKEIKPKDQMKKWMEM 480
    |||||||
Db 421 IGLSPDQEDNTEINFLKQALITVGLTPTMYLKKRMWVFKEIKPKDQMKKWMEM 480
QY 481 KREIVGVEPVPHDETYCPASLFHVSNDYSFIIRYTRTLXQFOFQALCOAKHEGPLH 540
    |||||||
Db 481 KREIVGVEPVPHDETYCPASLFHVSNDYSFIIRYTRTLXQFOFQALCOAKHEGPLH 540
QY 541 KCDISNSTEAGOKL 554
    |||||||
Db 541 KCDISNSTEAGOKL 554

```

RESULT 12

US-10-176-918-72

Sequence 72, Application US/10176918

Publication No. US20030027275A1

GENERAL INFORMATION:

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Desnoyers, Luc
: APPLICANT: Flivaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P330R1C382
: CURRENT APPLICATION NUMBER: US/10/176,918
: PRIORITY FILING DATE: 2002-06-20
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 72
: LENGTH: 555
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-176-918-72

```

```

Query Match 69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSSSSWLLSLVAVTAQSTIEQAKTFLLDKFNHEADLFYQSSLSANNTNITENVO 60
    |||||||
Db 1 MSSSSWLLSLVAVTAQSTIEQAKTFLLDKFNHEADLFYQSSLSANNTNITENVO 60
QY 61 MNMAGDKWSAFLEKQSTLAQNTPLQETIONLTVKLQALQOQSSVLSQDKSKRNTTL 120
    |||||||
Db 61 MNMAGDKWSAFLEKQSTLAQNTPLQETIONLTVKLQALQOQSSVLSQDKSKRNTTL 120
QY 121 NTMSTIYTGKVCNPDNPOECLELLEPGLINEIMANSLDYNERNLWAMESRSEVGKOLRPLY 180
    |||||||
Db 121 NTMSTIYTGKVCNPDNPOECLELLEPGLINEIMANSLDYNERNLWAMESRSEVGKOLRPLY 180
QY 181 EBYVVLKNEMARANHYEDYGDYWRGDEYVNGVDYDYSRGQLIEDVEHTFEELKPLYEHL 240
    |||||||
Db 181 EBYVVLKNEMARANHYEDYGDYWRGDEYVNGVDYDYSRGQLIEDVEHTFEELKPLYEHL 240
QY 241 HAYVRAKIMANAPSYISIPICLPAHLIGDMWGRFTNLXSLTVPFGOKPNIDVTAMDYQ 300
    |||||||
Db 241 HAYVRAKIMANAPSYISIPICLPAHLIGDMWGRFTNLXSLTVPFGOKPNIDVTAMDYQ 300
QY 301 AMDAORIFKEAEKFEFVSGLPNMTQGFWENSMLTDPGNVQKAVCHPTAMDLGKGFRIIM 360
    |||||||
Db 301 AMDAORIFKEAEKFEFVSGLPNMTQGFWENSMLTDPGNVQKAVCHPTAMDLGKGFRIIM 360
QY 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAPPELLRNGANEGHEAVGEIMSLSAATPKHLKS 420
    |||||||
Db 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAPPELLRNGANEGHEAVGEIMSLSAATPKHLKS 420
QY 421 IGLSPDQEDNTEINFLKQALITVGLTPTMYLKKRMWVFKEIKPKDQMKKWMEM 480
    |||||||
Db 421 IGLSPDQEDNTEINFLKQALITVGLTPTMYLKKRMWVFKEIKPKDQMKKWMEM 480
QY 481 KREIVGVEPVPHDETYCPASLFHVSNDYSFIIRYTRTLXQFOFQALCOAKHEGPLH 540
    |||||||
Db 481 KREIVGVEPVPHDETYCPASLFHVSNDYSFIIRYTRTLXQFOFQALCOAKHEGPLH 540
QY 541 KCDISNSTEAGOKL 554
    |||||||
Db 541 KCDISNSTEAGOKL 554

```

RESULT 13

US-10-176-921-72

Sequence 72, Application US/10176921

Publication No. US20030027276A1

GENERAL INFORMATION:

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Desnoyers, Luc
: APPLICANT: Flivaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P330R1C288
: CURRENT APPLICATION NUMBER: US/10/176,921
: PRIORITY FILING DATE: 2002-06-20
: Prior Application removed - See File Wrapper or Palm
: SEQ ID NO 72
: LENGTH: 555
: TYPE: PRT

```

ORGANISM: Homo Saplen
us-10-176-921-72

Query Match 69.4%; Score 2979; DB 9; Length 555;

Best Local Similarity 99.8%; Pred. No. 1.9e-239;

Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSSSSWLLSLVAVTAOSTIEBOAKTFPKFHNHEADLFYQSSLASMNNTNITEENVQ 60
DB 1 MSSSSWLLSLVAVTAOSTIEBOAKTFPKFHNHEADLFYQSSLASMNNTNITEENVQ 60
QY 61 MNNAAGDKWSAFLEKQSTLAQMPLOEIQNLTVKLOALQONGSSVLSSEKSKRLNTIL 120
DB 61 MNNAAGDKWSAFLEKQSTLAQMPLOEIQNLTVKLOALQONGSSVLSSEKSKRLNTIL 120
QY 121 NTMSTIYTGKVCNPDNPQECILLPEGLNEIMANSIDYNERLWAMESRSEVKQRLPLY 180
DB 121 NTMSTIYTGKVCNPDNPQECILLPEGLNEIMANSIDYNERLWAMESRSEVKQRLPLY 180
QY 181 EEEVVLKNEKMARANHEDYGDYWRGDEYNGVDYDYSRGQLIEDYHTEFEIKPLYEHL 240
DB 181 EEEVVLKNEKMARANHEDYGDYWRGDEYNGVDYDYSRGQLIEDYHTEFEIKPLYEHL 240
QY 241 HAYVRAKLMNAYPSYISPIGCLPAHLGDMGREFWNTLYSLTVPGOKPNIDVTDAVQ 300
DB 241 HAYVRAKLMNAYPSYISPIGCLPAHLGDMGREFWNTLYSLTVPGOKPNIDVTDAVQ 300
QY 301 AMDAQRIFKEAEKFFVSVGLPMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRIILM 360
DB 301 AMDAQRIFKEAEKFFVSVGLPMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRIILM 360
QY 361 CTKVTMDDEFLTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
DB 361 CTKVTMDDEFLTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
QY 421 IGLSPDFQEDNETEINFLLKQALITVGLPFTYMLEKRWVFKGEIIPKQDMKKWEM 480
DB 421 IGLSPDFQEDNETEINFLLKQALITVGLPFTYMLEKRWVFKGEIIPKQDMKKWEM 480
QY 481 KREIVGVEVPVPHDETCDPASLFHVSNDYSFIRYTRTLTYOFQFALCOAKHHEGPLH 540
DB 481 KREIVGVEVPVPHDETCDPASLFHVSNDYSFIRYTRTLTYOFQFALCOAKHHEGPLH 540
QY 541 KCDISNSTEAGOKL 554
DB 541 KCDISNSTEAGOKL 554

RESULT 14
US-10-137-865-72
; Sequence 72, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Mei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C154

```

CURRENT APPLICATION NUMBER: US/10/137,865
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO: 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Saplen
us-10-137-865-72

Query Match 69.4%; Score 2979; DB 9; Length 555;

Best Local Similarity 99.8%; Pred. No. 1.9e-239;

Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSSSSWLLSLVAVTAOSTIEBOAKTFPKFHNHEADLFYQSSLASMNNTNITEENVQ 60
DB 1 MSSSSWLLSLVAVTAOSTIEBOAKTFPKFHNHEADLFYQSSLASMNNTNITEENVQ 60
QY 61 MNNAAGDKWSAFLEKQSTLAQMPLOEIQNLTVKLOALQONGSSVLSSEKSKRLNTIL 120
DB 61 MNNAAGDKWSAFLEKQSTLAQMPLOEIQNLTVKLOALQONGSSVLSSEKSKRLNTIL 120
QY 121 NTMSTIYTGKVCNPDNPQECILLPEGLNEIMANSIDYNERLWAMESRSEVKQRLPLY 180
DB 121 NTMSTIYTGKVCNPDNPQECILLPEGLNEIMANSIDYNERLWAMESRSEVKQRLPLY 180
QY 181 EEEVVLKNEKMARANHEDYGDYWRGDEYNGVDYDYSRGQLIEDYHTEFEIKPLYEHL 240
DB 181 EEEVVLKNEKMARANHEDYGDYWRGDEYNGVDYDYSRGQLIEDYHTEFEIKPLYEHL 240
QY 241 HAYVRAKLMNAYPSYISPIGCLPAHLGDMGREFWNTLYSLTVPGOKPNIDVTDAVQ 300
DB 241 HAYVRAKLMNAYPSYISPIGCLPAHLGDMGREFWNTLYSLTVPGOKPNIDVTDAVQ 300
QY 301 AMDAQRIFKEAEKFFVSVGLPMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRIILM 360
DB 301 AMDAQRIFKEAEKFFVSVGLPMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRIILM 360
QY 361 CTKVTMDDEFLTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
DB 361 CTKVTMDDEFLTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
QY 421 IGLSPDFQEDNETEINFLLKQALITVGLPFTYMLEKRWVFKGEIIPKQDMKKWEM 480
DB 421 IGLSPDFQEDNETEINFLLKQALITVGLPFTYMLEKRWVFKGEIIPKQDMKKWEM 480
QY 481 KREIVGVEVPVPHDETCDPASLFHVSNDYSFIRYTRTLTYOFQFALCOAKHHEGPLH 540
DB 481 KREIVGVEVPVPHDETCDPASLFHVSNDYSFIRYTRTLTYOFQFALCOAKHHEGPLH 540
QY 541 KCDISNSTEAGOKL 554
DB 541 KCDISNSTEAGOKL 554

RESULT 15
US-10-140-474-72
; Sequence 72, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Mei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.

```

```

; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C162
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 72
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-474-72

Query Match          69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSWLLSLVAVTAOSTIEEOKTFLDKFNHEAEDLFYQSSLASMYNTNITEENVQ 60
Db 1 MSSSSWLLSLVAVTAOSTIEEOKTFLDKFNHEAEDLFYQSSLASMYNTNITEENVQ 60
QY 61 MNNAAGDKWSAFLEKOSTLAQMPLOEIQLTVKLOLQALQONGSSVLSDESKRLNTIL 120
Db 61 MNNAAGDKWSAFLEKOSTLAQMPLOEIQLTVKLOLQALQONGSSVLSDESKRLNTIL 120
QY 121 NTMSTIYSGKVCNPNPQECLELLEPGNEIMANSIDYNERLWAMWSRSEVKGOLRPLY 180
Db 121 NTMSTIYSGKVCNPNPQECLELLEPGNEIMANSIDYNERLWAMWSRSEVKGOLRPLY 180
QY 181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRQGLIEDVEHTEFEIKPLYEHL 240
Db 181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRQGLIEDVEHTEFEIKPLYEHL 240
QY 241 HAYYRAKIMNAYPSYISPIGCLPAHLGDMWGFMTNLYSLTPFGQKPNIDVTDAMVDQ 300
Db 241 HAYYRAKIMNAYPSYISPIGCLPAHLGDMWGFMTNLYSLTPFGQKPNIDVTDAMVDQ 300
QY 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWNSMLTDPGNVQKAVCHPTAMDLGKGFRLIM 360
Db 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWNSMLTDPGNVQKAVCHPTAMDLGKGFRLIM 360
QY 361 CTKYTMDDDLTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
Db 361 CTKYTMDDDLTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
QY 421 IGLSPFOEDNTEINFLKQALTYGTLPTTYMLEKMRMVFKEGIPKQMMKKWEM 480
Db 421 IGLSPFOEDNTEINFLKQALTYGTLPTTYMLEKMRMVFKEGIPKQMMKKWEM 480
QY 481 KREIVGVVEPVPHDETYCDPASLFFHVSNDYSFIRYTRTLVYQFOQALCOAKHHEGRLH 540
Db 481 KREIVGVVEPVPHDETYCDPASLFFHVSNDYSFIRYTRTLVYQFOQALCOAKHHEGRLH 540
QY 541 KCDISNSTEAGQKL 554
Db 541 KCDISNSTEAGQKL 554

RESULT 16
US-10-142-431-72
; Sequence 72, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang

```

```

; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C251
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 72
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-431-72

Query Match          69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSWLLSLVAVTAOSTIEEOKTFLDKFNHEAEDLFYQSSLASMYNTNITEENVQ 60
Db 1 MSSSSWLLSLVAVTAOSTIEEOKTFLDKFNHEAEDLFYQSSLASMYNTNITEENVQ 60
QY 61 MNNAAGDKWSAFLEKOSTLAQMPLOEIQLTVKLOLQALQONGSSVLSDESKRLNTIL 120
Db 61 MNNAAGDKWSAFLEKOSTLAQMPLOEIQLTVKLOLQALQONGSSVLSDESKRLNTIL 120
QY 121 NTMSTIYSGKVCNPNPQECLELLEPGNEIMANSIDYNERLWAMWSRSEVKGOLRPLY 180
Db 121 NTMSTIYSGKVCNPNPQECLELLEPGNEIMANSIDYNERLWAMWSRSEVKGOLRPLY 180
QY 181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRQGLIEDVEHTEFEIKPLYEHL 240
Db 181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRQGLIEDVEHTEFEIKPLYEHL 240
QY 241 HAYYRAKIMNAYPSYISPIGCLPAHLGDMWGFMTNLYSLTPFGQKPNIDVTDAMVDQ 300
Db 241 HAYYRAKIMNAYPSYISPIGCLPAHLGDMWGFMTNLYSLTPFGQKPNIDVTDAMVDQ 300
QY 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWNSMLTDPGNVQKAVCHPTAMDLGKGFRLIM 360
Db 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWNSMLTDPGNVQKAVCHPTAMDLGKGFRLIM 360
QY 361 CTKYTMDDDLTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
Db 361 CTKYTMDDDLTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
QY 421 IGLSPFOEDNTEINFLKQALTYGTLPTTYMLEKMRMVFKEGIPKQMMKKWEM 480
Db 421 IGLSPFOEDNTEINFLKQALTYGTLPTTYMLEKMRMVFKEGIPKQMMKKWEM 480
QY 481 KREIVGVVEPVPHDETYCDPASLFFHVSNDYSFIRYTRTLVYQFOQALCOAKHHEGRLH 540
Db 481 KREIVGVVEPVPHDETYCDPASLFFHVSNDYSFIRYTRTLVYQFOQALCOAKHHEGRLH 540
QY 541 KCDISNSTEAGQKL 554
Db 541 KCDISNSTEAGQKL 554

RESULT 17
US-10-143-114-72
; Sequence 72, Application US/10143114
; Publication No. US20030036180A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C211
; CURRENT APPLICATION NUMBER: US/10/143,114
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 72
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-143-114-72

```

```

Query Match          69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSSSSMLLSLVAVTAOSTIEEOAKTEFLDKFNHEADLFYQSSLSAWNTNTTEENVO 60
DB 1 MSSSSMLLSLVAVTAOSTIEEOAKTEFLDKFNHEADLFYQSSLSAWNTNTTEENVO 60
QY 61 MNNAAGDKWSAFLKEOSTLAQMPPLQEIQNLTVKQLQALQOGSSVLSDEKSKRLNTIL 120
DB 61 MNNAAGDKWSAFLKEOSTLAQMPPLQEIQNLTVKQLQALQOGSSVLSDEKSKRLNTIL 120
QY 121 NTMSTIYSTGKVCNPDPQECLEPGLNEIMANSLDYNERLWAMESRSEVKGOLRPLX 180
DB 121 NTMSTIYSTGKVCNPDPQECLEPGLNEIMANSLDYNERLWAMESRSEVKGOLRPLX 180
QY 181 EEEYVLKNEMARANHYEDYGDYWRGDEYVNGYDYSRQGLIEVHTPEEIKPLYEHL 240
DB 181 EEEYVLKNEMARANHYEDYGDYWRGDEYVNGYDYSRQGLIEVHTPEEIKPLYEHL 240
QY 241 HAYYRAKLMAVPSYISPIGCLPAHLIGDMGREFWNLVSLTVPGOKPNIDVTAMVDO 300
DB 241 HAYYRAKLMAVPSYISPIGCLPAHLIGDMGREFWNLVSLTVPGOKPNIDVTAMVDO 300
QY 301 AMDAORIFKEAEKFFVSVGLPNNTOGFWENSMITDPGNVOKAVCHPTAMDLGKDFRILM 360
DB 301 AMDAORIFKEAEKFFVSVGLPNNTOGFWENSMITDPGNVOKAVCHPTAMDLGKDFRILM 360
QY 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAOPFLIRNGANEGHEAVGEIMSLSAATPRHLKS 420
DB 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAOPFLIRNGANEGHEAVGEIMSLSAATPRHLKS 420
QY 421 IGLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWMMVFGEIPKQDMKKWEM 480
DB 421 IGLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWMMVFGEIPKQDMKKWEM 480
QY 481 KREIVGVEVPVPHDETYCPASLFHVSNDYSFIRYRTLYQFQOEALQCAAKHGGPLH 540
DB 481 KREIVGVEVPVPHDETYCPASLFHVSNDYSFIRYRTLYQFQOEALQCAAKHGGPLH 540
QY 541 KCDISNSTAGOKL 554
DB 541 KCDISNSTAGOKL 554

```

```

DB 541 KCDISNSTAGOKL 554

```

```

RESULT 18
US-10-140-002-72
; Sequence 72, Application US/10140002
; Publication No. US20030037623A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C59
; CURRENT APPLICATION NUMBER: US/10/140,002
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 72
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-002-72

```

```

Query Match          69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSSSSMLLSLVAVTAOSTIEEOAKTEFLDKFNHEADLFYQSSLSAWNTNTTEENVO 60
DB 1 MSSSSMLLSLVAVTAOSTIEEOAKTEFLDKFNHEADLFYQSSLSAWNTNTTEENVO 60
QY 61 MNNAAGDKWSAFLKEOSTLAQMPPLQEIQNLTVKQLQALQOGSSVLSDEKSKRLNTIL 120
DB 61 MNNAAGDKWSAFLKEOSTLAQMPPLQEIQNLTVKQLQALQOGSSVLSDEKSKRLNTIL 120
QY 121 NTMSTIYSTGKVCNPDPQECLEPGLNEIMANSLDYNERLWAMESRSEVKGOLRPLX 180
DB 121 NTMSTIYSTGKVCNPDPQECLEPGLNEIMANSLDYNERLWAMESRSEVKGOLRPLX 180
QY 181 EEEYVLKNEMARANHYEDYGDYWRGDEYVNGYDYSRQGLIEVHTPEEIKPLYEHL 240
DB 181 EEEYVLKNEMARANHYEDYGDYWRGDEYVNGYDYSRQGLIEVHTPEEIKPLYEHL 240
QY 241 HAYYRAKLMAVPSYISPIGCLPAHLIGDMGREFWNLVSLTVPGOKPNIDVTAMVDO 300
DB 241 HAYYRAKLMAVPSYISPIGCLPAHLIGDMGREFWNLVSLTVPGOKPNIDVTAMVDO 300
QY 301 AMDAORIFKEAEKFFVSVGLPNNTOGFWENSMITDPGNVOKAVCHPTAMDLGKDFRILM 360
DB 301 AMDAORIFKEAEKFFVSVGLPNNTOGFWENSMITDPGNVOKAVCHPTAMDLGKDFRILM 360
QY 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAOPFLIRNGANEGHEAVGEIMSLSAATPRHLKS 420
DB 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAOPFLIRNGANEGHEAVGEIMSLSAATPRHLKS 420
QY 421 IGLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWMMVFGEIPKQDMKKWEM 480
DB 421 IGLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWMMVFGEIPKQDMKKWEM 480

```

QY 481 KREIVGVPEVPHDETCDBASLFHVSNDYSFIRYTRTLXQFOFOEALCOAAKHGEPH 540
 DB 481 KREIVGVPEVPHDETCDBASLFHVSNDYSFIRYTRTLXQFOFOEALCOAAKHGEPH 540
 QY 541 KCDISNSTEAGOKL 554
 DB 541 KCDISNSTEAGOKL 554

RESULT 19

US-10-142-419-72
 ; Sequence 72, Application US/10142419
 ; Publication No. US20030044945A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3330R1C244
 ; CURRENT APPLICATION NUMBER: US/10/142,419
 ; CURRENT FILING DATE: 2002-05-10
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 72
 ; LENGTH: 555
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-142-419-72

Query Match 69.4%; Score 2979; DB %: Length 555;
 Best Local Similarity 99.8%; Pred. No. 1.9e-239;
 Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSWLLSLVAVTAOSTIEEQAKTEFLDFNHEADLFYQSSSLASWNTNTITEENVQ 60
 DB 1 MSSSSWLLSLVAVTAOSTIEEQAKTEFLDFNHEADLFYQSSSLASWNTNTITEENVQ 60
 QY 61 MNNAGDKWSAFLKEQSTLAQMPLOETIQLTVKLOLQALQONSSVLSSEDKSKRLNTIL 120
 DB 61 MNNAGDKWSAFLKEQSTLAQMPLOETIQLTVKLOLQALQONSSVLSSEDKSKRLNTIL 120
 QY 121 NTMSTIYSTGVCNPDNQECLLEPGINETMANSLDYNERLWAMESRSRSEVGKOLRPLY 180
 DB 121 NTMSTIYSTGVCNPDNQECLLEPGINETMANSLDYNERLWAMESRSRSEVGKOLRPLY 180
 QY 181 EYVVLKEMARAHYEDYGYWRGDEYVNGVDYDSRGQLIEDVHTTEEEKPLYEHL 240
 DB 181 EYVVLKEMARAHYEDYGYWRGDEYVNGVDYDSRGQLIEDVHTTEEEKPLYEHL 240
 QY 241 HAYVRAKLMAVPSYISPIGCLPAHLIGDMWGRFTWNLVSLTFVFGOKPNTDVTDAVQ 300
 DB 241 HAYVRAKLMAVPSYISPIGCLPAHLIGDMWGRFTWNLVSLTFVFGOKPNTDVTDAVQ 300
 QY 301 AMDAQRIFKEAEKFFVSVGLPNTQGFWENSMLTDPGNVOKAVCHPTAMDLGKDFRILM 360
 DB 301 AMDAQRIFKEAEKFFVSVGLPNTQGFWENSMLTDPGNVOKAVCHPTAMDLGKDFRILM 360

QY 361 CTKVTMDPELTAHHEMHTQYDMAAQAOPFLLRNCGFHEANGELMSLSAATPKHLKS 420
 DB 361 CTKVTMDPELTAHHEMHTQYDMAAQAOPFLLRNCGFHEANGELMSLSAATPKHLKS 420
 QY 421 IGLSPDFQEDNETELNFKALITVGTLPFTVLEKRWNVFPGELPKOAMKRWEM 480
 DB 421 IGLSPDFQEDNETELNFKALITVGTLPFTVLEKRWNVFPGELPKOAMKRWEM 480
 QY 481 KREIVGVPEVPHDETCDBASLFHVSNDYSFIRYTRTLXQFOFOEALCOAAKHGEPH 540
 DB 481 KREIVGVPEVPHDETCDBASLFHVSNDYSFIRYTRTLXQFOFOEALCOAAKHGEPH 540
 QY 541 KCDISNSTEAGOKL 554
 DB 541 KCDISNSTEAGOKL 554

RESULT 20

US-09-969-384-23
 ; Sequence 23, Application US/09969384
 ; Publication No. US20020192749A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Moore, et al.
 ; TITLE OF INVENTION: Human Gene Polynucleotides, polypeptides, and Antibodies
 ; FILE REFERENCE: PM055P1
 ; CURRENT APPLICATION NUMBER: US/09/969,384
 ; PRIOR FILING DATE: 2001-10-03
 ; PRIOR APPLICATION NUMBER: PCT/US01/10542
 ; PRIOR FILING DATE: 2001-04-02
 ; PRIOR APPLICATION NUMBER: 60/236,384
 ; PRIOR FILING DATE: 2000-09-29
 ; PRIOR APPLICATION NUMBER: 60/194,118
 ; PRIOR FILING DATE: 2000-04-03
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: Patentln Ver. 2.0
 ; SEQ ID NO 23
 ; LENGTH: 261
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-969-384-23

Query Match 31.7%; Score 1359; DB %: Length 261;
 Best Local Similarity 99.6%; Pred. No. 3.1e-105;
 Matches 252; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 62 MNNAGDKWSAFLKEQSTLAQMPLOETIQLTVKLOLQALQONSSVLSSEDKSKRLNTIL 121
 DB 1 MNNAGDKWSAFLKEQSTLAQMPLOETIQLTVKLOLQALQONSSVLSSEDKSKRLNTIL 121
 QY 122 TMSTIYSTGVCNPDNQECLLEPGINETMANSLDYNERLWAMESRSRSEVGKOLRPLY 181
 DB 122 TMSTIYSTGVCNPDNQECLLEPGINETMANSLDYNERLWAMESRSRSEVGKOLRPLY 181
 QY 182 EYVVLKEMARAHYEDYGYWRGDEYVNGVDYDSRGQLIEDVHTTEEEKPLYEHL 241
 DB 182 EYVVLKEMARAHYEDYGYWRGDEYVNGVDYDSRGQLIEDVHTTEEEKPLYEHL 241
 QY 242 AYVRAKLMAVPSYISPIGCLPAHLIGDMWGRFTWNLVSLTFVFGOKPNTDVTDAVQ 301
 DB 242 AYVRAKLMAVPSYISPIGCLPAHLIGDMWGRFTWNLVSLTFVFGOKPNTDVTDAVQ 301
 QY 302 WDAQRIFKEAEK 314
 DB 241 WDAQRIFKEAEK 253

Search completed: March 13, 2003, 17:02:41
 Job time : 25 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 7, 2003, 06:35:39 ; Search time 4862 Seconds
(without alignments)
4818.544 Million cell updates/sec

Title: US-09-978-385-2

Perfect score: 4291
Sequence: 1 MSSSSWLLSLVAVTAQST.....ISKENNPQNTDVTQTSF 805

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xlp
-MODE=frame+pn.model -US09978385/runat.28022003.104716.1342/app.query.fasta.1.967
-Q/cgn2.1/USPTO.spool/US09978385 -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=20 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09978385 @cgn1.1.3745 @runat.28022003.104716.1342 -NCPU=6 -ICPU=3
-WNO.XLPXY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WNO.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_to:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4291	100.0	2415	6 AR135178	AR135178 Sequence
2	4291	100.0	2415	6 AX418984	AX418984 Sequence
3	4291	100.0	2415	6 E43987	E43987 ACF-analog
4	4291	100.0	2418	6 E39033	E39033 MPROT15 pol
5	4291	100.0	2599	6 E43988	E43988 ACF-analog
6	4291	100.0	3325	6 AF291820	AF291820 Homo sapi
7	4291	100.0	3334	6 AX047758	AX047758 Sequence
8	4291	100.0	3396	6 AX418982	AX418982 Sequence
9	4291	100.0	3396	6 AX418982	AX418982 Sequence
10	4291	100.0	3396	6 AX431513	AX431513 Sequence
11	4291	100.0	3405	6 AF241254	AF241254 Homo sapi
12	4288	99.9	3341	9 AB046569	AB046569 Homo sapi
13	4281	99.8	3341	6 AR135177	AR135177 Sequence
14	4142	96.5	3732	6 AX463938	AX463938 Sequence
15	3740.5	87.2	2262	6 E39034	E39034 MPROT15 pol
16	3579	83.4	2638	6 AX047762	AX047762 Sequence
17	3579	83.4	2739	10 BC026801	BC026801 Mus muscu
18	3562	83.0	2760	10 AB053181	AB053181 Mus muscu
19	3561	83.0	2638	6 AX047765	AX047765 Sequence
20	3509	81.8	2415	6 AX047764	AX047764 Sequence
21	2904	67.7	2415	6 BC034367	BC034367 Mus muscu
22	2167	50.5	1993	10 BC034367	BC034367 Mus muscu
23	1347.5	31.4	3681	9 BC034367	BC034367 Mus muscu
24	1347	31.4	2473	9 HSMACEA	HSMACEA
25	1344	31.3	2477	6 HSMACEA	HSMACEA
26	1344	31.3	2477	6 HSMACEA	HSMACEA
27	1344	31.3	2478	6 AR037213	AR037213 Sequence
28	1344	31.3	2478	6 AR137383	AR137383 Sequence
29	1337	31.2	4020	6 AR137383	AR137383 Sequence
30	1337	31.2	4020	6 HUMACB	HUMACB
31	1337	31.2	4022	6 AX429555	AX429555 Sequence
32	1337	31.2	4024	6 A00914	A00914 H. sapiens g
33	1334	31.1	2418	10 MUSACB	MUSACB
34	1334	31.1	3813	10 MUSACB	MUSACB
35	1334	31.1	3939	6 AX147531	AX147531 Sequence
36	1334	31.1	4563	6 AX147503	AX147503 Sequence
37	1334	31.1	4694	10 MUSACB	MUSACB
38	1332	30.6	4050	5 CHKACEI	CHKACEI
39	1312	30.5	4014	10 AF201331	AF201331 Rattus norv
40	1310	30.5	4014	10 AF201332	AF201332 Rattus norv
41	1310	30.5	4014	10 RND03708	RND03708 Rattus norv
42	1310	30.5	4014	10 RND03708	RND03708 Rattus norv
43	1310	30.5	4142	10 RND03734	RND03734 Rattus norv
44	1303	30.4	2490	9 HUMTACB	HUMTACB
45	1283.5	29.9	4803	4 OCANCOE	OCANCOE

RESULT 1

ALIGNMENTS

AR135178
 LOCUS AR135178 2415 bp DNA linear PAT 16-MAY-2001
 DEFINITION Sequence 3 from patent US 6194556.
 ACCESSION AR135178
 VERSION AR135178.1 GI:14124083
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclassified.
 1 (bases 1 to 2415)
 AUTHORS Acton, S. Laurene, and Robison, K. Earl.
 TITLE Angiotensin converting enzyme homolog and therapeutic and
 diagnostic uses therefor
 JOURNAL Patent: US 6194556-A 3 27-FEB-2001;
 FEATURES Location/Qualifiers
 source 1..2415
 BASE COUNT 743 a 483 c 555 g 634 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2415
 Score: 4291.00 Matches: 805
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 US-09-978-385-2 (1-805) x AR135178 (1-2415)
 QY 1 MetSerSerSerSerTriPLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
 Db 1 ATGTCAAGCTCTCTCGCTGCTCTCTCAAGCTGTGCTTAACTGCTGCTGCTCAGCCACC 60
 QY 21 IleGluGluGlnAlaValThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
 Db 61 ATTAGGAACAGGCCAAGACATTTTGGACAGTTTAAACAGCAACCCAGAACCTGTTTC 120
 QY 41 TyrGlnSerSerLeuAlaSerTriPAsnTyrAsnThrAsnIleThrGluGluAsnValGln 60
 Db 121 TATCAAGTTCACCTGCTTCTTGAATTTATACCAAAATTTATCTGAAGAAGATGTCCA 180
 QY 61 AsnMetAsnAsnAlaGlyAspLysTyrPserAlaPheLeuGluGlnSerThrLeuAla 80
 Db 181 AACATGATATATGCTGGGCAAAATGTCTGCTTTTAAAGAACAGTCCACTTGGC 240
 QY 81 GlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuAlaLeu 100
 Db 241 CAATGTATCCTACTACAGAAATTCAGAAATTCACAGTCAAGCTTCAGCTGAGGCTCTT 300
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrLeu 120
 Db 301 CACCAAAATGGGCTCTCACTGCTGCAGAACCAAGAACCAAGCGTTTAAACCAATTTCT 360
 QY 121 AsnThrMetSerThrIleTyrSerThrArgLysValLysAsnProAspAsnProGlnGlu 140
 Db 361 AATACATATAGACCACTTCAAGAAATTCAGAAATTTGTAAACCAAGATATCCACAGAA 420
 QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
 Db 421 TCGTATATCTTGAACCAAGGTTTGAATTAATATGCAAAACATTTTGAAGTAAATGAG 480
 QY 161 ArgLeuThrAlaTyrGluSerTyrPargSerGluValGlyLysGlnLeuArgProLeuTyr 180
 Db 481 AGGCTCTGGGCTTGGAAAGCTGGAGATCTGAGGCGCAAGCAGCTGAGCCATTATAT 540
 QY 181 GluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200
 Db 541 GAAAGATATGTGCTTAAATAAAGAGTGGCAAGACCAATCATATATGAGGACTATGCG 600
 QY 201 AspTyrTTPargGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
 Db 601 GATTATTTGAGAGAGACTATGAAATGAGGCTAGATGCTATGACTACGCCCGGC 660
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240
 Db 661 CAGTTATTTGAAGATGTGGAACATACCTTTGAAGAGATTAACCTTTATGACATCTT 720
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
 Db 721 CATGCTATGTAGAGGCCAAAGTTGATGAAATGCAATTCCTTCTATATCAGTCAATTTGA 780
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheTyrPThrAsnLeuTyrSer 280
 Db 781 TGCTTCCCTGCTCATATTTGCTGTGATATGAGGGGTATATTTTGACAAATCTGTACTCT 840
 QY 281 LeuThrValProPheGlyGluLysProAsnIleAspValThrAspAlaMetValAspGln 300
 Db 841 TTGACAGTTCCTTTGGACAGAAACCAATCATATGTTACTGATCATGAGTGGAGCCAG 900
 QY 301 AlaTyrPaspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320
 Db 901 GCCTGGATGACAGAGAAATTTCAAGAGAGCCAGAAAGTTCTTTTATCTGTTGCTT 960
 QY 321 ProAsnMetThrGlnGlyPheTyrGluAsnSerMetLeuThrAspProGlyAsnValGln 340
 Db 961 CCTAATATGACTCAAGGATTTCTGGAAATTCATGCTATACGAGACCAAGAAATGTTTCA 1020
 QY 341 LysAlaValCysHisProThrAlaTyrPaspLeuGlyLysGlyAspPheArgIleLeuMet 360
 Db 1021 AAAGCAGCTCTCCATCCACAGCTTGGAGCTGGGAGAGGCCACTTCAGAGATCTTATG 1080
 QY 361 CysThrLysValIlePheMetAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380
 Db 1081 TGCCAAAGGAGAACAAATGAGCGACTTCTGACAGCTCATCATGAGATGGGCAATTCAG 1140
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyValAsnGluLysPhe 400
 Db 1141 TATATATGCAATATGCTGACAAACCTTTCTCTTAAGAAAGAGCTTAATGAAAGATTC 1200
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 Db 1201 CATGACCTTGTGGGAAATCATGCTCTTCTGCGCACACCTTAACCATTTAAATCC 1260
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440
 Db 1261 ATTGCTCTTGTGACCCGATTTTCAAGAAAGCAATGAAACGAAATTAACCTTCTGCTC 1320
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTyrParg 460
 Db 1321 AAACAGACACTCAGATTTGTTGGACTGCTGCAATTTACTTACATGTTAAGAAATGAGG 1380
 QY 461 TrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTyrPargLeuMet 480
 Db 1381 TGGATGCTTTTAAAGGGGAAATTTCCCAAAAGCAGTGGAGTGAAGAGTGGGAGATG 1440
 QY 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500
 Db 1441 AAGCGAGAGATGTTGGGTGGTGGAACTGTGCCCATGATGAACATATCTGTGACCC 1500
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520
 Db 1501 GCATCTGCTTCCATGTTTCTTAATGATTAATCAATTCATTTGATATTCACAAAGACCTT 1560
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaAlaLysHisGluGlyProLeuHis 540
 Db 1561 TACCAATTTCCAGTTTCAAGAGCACTTGTCAAGACCTTAACATGAGAGCCCTTGTGAC 1620
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560
 Db 1621 AAATGTGACATCTCAACTTACAGAACTGAGACAGAACTGTTCAATATGCTGAGGCTT 1680
 QY 561 GlyLysSerGluProTyrPThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsn 580
 Db 1681 GGAATATCAGAACCTTGACCTTACGATTTGGAAATGTTGTAGGAGCAAGAAATGAT 1740

```

QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPLeuLysAspGlnAsnLys 600
Db 1741 GTAAGGCCACTGCTCACTACTTGTAGCCCTATTATTTACCTGGCTGAAGACAGAACAG 1800
QY 601 AsnSerPheValGlyTyrSerThrAspTrpSerProTyrAlaAspGlnSerLleLysVal 620
Db 1801 AATTCTTTTGGGATGGAGTACCGACTGGATCCATATGAGACCAAGCATCAAAATG 1860
QY 621 ArgLleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPasnAspAsnGlnMet 640
Db 1861 AGGATAAGCTTAATAATCAGCTCTTGGAGATTAAGCATATGATGAGACGAATGAATG 1920
QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetLargGlnTyrPheLeuLysValLysAsn 660
Db 1921 TACCTGTCGATCATCTGTTGATATGCTATGAGCCAGTCTTTTAAAGTAAAAAT 1980
QY 661 GlnMetLleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgLleSer 680
Db 1981 CAGATGATCTTTTGGGAGAGAGATGTGCGAGTGGCTTAATTTGAAACCAAGATCTCC 2040
QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspLleLleProArgThrGluVal 700
Db 2041 TTATATTTCTTGTGCACAGCAGCTAAATAATGTCTGATATCATCTCTAGACTGAAGTT 2100
QY 701 GluLysAlaIleArgMetSerArgSerArgLleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2101 GAAAGGCCATCAGAGATGTCGCCGAGCCGTATCATGATGCTTCCGTGTAATGACAC 2160
QY 721 SerLeuGluPheLeuGlyLleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2161 AGCCTAAGTCTTCTGGGAGTACAGCCCAACACTTGACCTCTTAACCGCCCTGTTTCC 2220
QY 741 IleThrLeuLleValPheGlyValAlaMetGlyValLleValGlyLleValLleLeu 760
Db 2221 ATATGGCTGATGTTGTTGGAGTGTGATGGAGTGAATGAGTGGCATTTGATCTC 2280
QY 761 IlePheThrGlyLleArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnPro 780
Db 2281 ATCTTCACCTGGATCAGAGATCGGAAAGAAAAATTAACCAAGACTGAGAAAAATCCT 2340
QY 781 TyrAlaSerLleAspLleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAsp 800
Db 2341 TATCCTTCATCGATATTAGCAAAAGAAATATATCCAGATTTCCAAAACACTGATGAT 2400
QY 801 ValGlnThrSerPhe 805
Db 2401 GTTCAGACTGCTCTT 2415

RESULT 2
AX418984
LOCUS AX418984 2415 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 3 from Patent WO0212471.
ACCESSION AX418984
VERSION AX418984.1 GI:21523783
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Action, S., Robison, K.E. and Hsieh, F.Y.
Angiotensin converting enzyme homolog
Patent: WO 0212471-A 3 14-FEB-2002;
Millennium Pharmaceuticals, Inc. (US)

FEATURES
source
1..2415
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 743 a 483 c 555 g 634 t
ORIGIN
Alignment Scores: 0 Length: 2415
Pred. No.:

```

```

Score: 4291.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 6 Gaps: 0

US-09-978-385-2 (1-805) x AX418984 (1-2415)

QY 1 MetSerSerSerSerThrPLeuLeuLeuSerLeuValAlaValThrAlaIleGlnSerThr 20
Db 1 ATGTCAAGCTCTTCTCGGCTCTTCTCAGCCTTGTTGCTGTACTCTGCTGCTGCTGCTGCT 60
QY 21 IleGluGluGlnLalysThrPheLeuAspLysPheAsnHisGluLalagLunsPhePhe 40
Db 61 ATTGAGACAGGCCAAGACATTTTGGACAAATTTTACCACGAGCCGAAAGCCGTTC 120
QY 41 TyrGlnSerSerLeuAlaSerTyrPasnTyrAsnThrAsnLleThrGluGluAsnValGln 60
Db 121 TATCAAAAGTTCACCTTGCTGCTTGTGATTTAAACACCAATTTATGAAAGAAATGCTCCA 180
QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
Db 181 AACATGAATATATGCTGGGCAAAATGCTGCTCTTTTAAAGGAACAGTCCACACTTGGC 240
QY 81 GlnMetTyrProLeuGlnGluLleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
Db 241 CAATATGATTCACCTTACAGAAATTCAGAAATTCACAGTCAAGCTTCAGCTGAGCTCTT 300
QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrLleLeu 120
Db 301 CACCAAAATGGGCTTCAGTGGCTGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 360
QY 121 AsnThrMetSerThrLleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
Db 361 AATACATATGAGCAACCTTCACAGTCTGAGAAAGTTTGAACCCAGATATCCAAAGAA 420
QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGlnLleMetAlaAsnSerLeuAspTyrAsnGlu 160
Db 421 TGGTATTTACTTGAACACAGTGTGAATGAATTAATGGCAACAGTTTACACTACATGAG 480
QY 161 ArgLeuThrAlaIleTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyr 180
Db 481 AGCCTCTGGCTTGGGAAAGCTGAGATCTGAGTGGCAACAGCTGAGGCAATATAT 540
QY 181 GluGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAspTyrGly 200
Db 541 GAAAGATATGCTGCTTGAATAATGAGATGGCAAGCAAAATCATATGAGAGCTATGAGG 600
QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
Db 601 GATTATTTGGAGAGAGACTATGAAATGAGGATGAGTATGCTATGACTACAGCCGCGC 660
QY 221 GlnLeuLleGluAspValGlnHisThrPheGluGluLleLysProLeuTyrGluHisLeu 240
Db 661 CAGTTGATTTGAAGATGTGAAACATACCTTTGAAGAGATTAAACCATATATGAAACATCTT 720
QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrLleSerProLleGly 260
Db 721 CATGCTATATGAGAGGCAAAAGTGAATGATGCTATCTCTATATGAGTCCAAATTTGGA 780
QY 261 CysLeuProAlaHisLeuLeuLysAspMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280
Db 781 TGCCCTCCCTGCACTTCTGCTGATGTGTGGGCTAGATTTTGGACAAATCTGTACTCT 840
QY 281 LeuThrValProPheGlyGluLysProAsnLleAspValThrAspAlaMetValAspGln 300
Db 841 TTGAAGATTCCTTTGGGACAGAAACCAACATAGATGTACTGATGCAATGTGTGAGCAG 900
QY 301 AlaTrpAspAlaGlnArgLlePheLysGluLalagLunsPhePheValSerValGlyLeu 320
Db 901 GCTGGGATGCACAGAGAAATATTCAGAGAGCCGAGAAATCTTTGATGATGCTGCTT 960
QY 321 ProAsnMetThrGlnGlyPheThrGluAsnSerMetLeuThrAspProGlyAsnValGln 340

```

```

|||||
Db CCTAATATGACTCAAGGATTCGGGAAATTCATGCTTACCGACGACCGAAATGTTCCG 1020
Qy 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1021 AAGGACGCTGCTCCATCCACAGCTTGGGACCTGGGAAAGGCGACTTCAGATCTTATG 1080
Qy 361 CysThrLysValIleMetAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380
Db 1081 TGCACAAGGTGACAAATGAGCGACTTCCTGACAGCTCATCATGATGATGGGCAATCCAG 1140
Qy 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAspGlyLysGlnGlyPhe 400
Db 1141 TATGATATGGCATATGCTGCAACCTTTCTCTAAGAAATGAGCTAATGAAAGATTTC 1200
Qy 401 HisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
Db 1201 CATGACGCTTGGGAAATCATGTCATCTTTCGACGCCACACCTTAAGCATTTAAATCC 1260
Qy 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnIleLeu 440
Db 1261 ATTGCTCTCTGTCACCCGATTTTCAAGAAACAAATGAAACAAATTAACCTTCGCTC 1320
Qy 441 LysGlnAlaLeuThrIleValIleGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460
Db 1321 AACCAAGCACTCACGATTTGGGACTCTGCCATTTACTTACATGTTAGAAAGTGGAGG 1380
Qy 461 TrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTrpArgIleMet 480
Db 1381 TGGATGCTCTTAAAGGGGAAATTCCTCAAGACAGATGATGAAAGTGGTGGGAGATG 1440
Qy 481 LysArgGluIleValIleGlyValIleGluProValProHisAspGluThrTyrCysAspPro 500
Db 1441 AAGGAGAGATAGTGGGGGTGGGAGACCTGGCCCATATGAAATCTGTGACCCC 1500
Qy 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrpArgThrLeu 520
Db 1501 GCATCTCTGTTCCATCTTTCTTAAGATTAATCTCATCTTATGATATTCACAAAGACCTT 1560
Qy 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis 540
Db 1561 TACCAATTCCTCAAGGAAAGCACTTTGTCAGACACCTTAACATGTAAGGCCCTCGCAC 1620
Qy 541 LysCysAspIleSerAsnSerThrGluAlaGlyLysIleLeuPheAsnMetLeuArgLeu 560
Db 1621 AAAAGTGCATCTCAACTCTACAGAACTGGACAGAACTGTTCAATATAGCTGAGCCTT 1680
Qy 561 GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValIleGlyAlaLysAsnMetAsn 580
Db 1681 GGAAATCAGAAACCTCGACCTGACCTTGCAAAATGTTTGAGAGACAAAGAACATCAT 1740
Qy 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600
Db 1741 GTAAGGCCACACTCACTTACTTATGAGCCCTTATTTACTGCTGGGAAAGCCACAAAG 1800
Qy 601 AsnSerPheValIleTyrPheSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620
Db 1801 AATCTTTTGGGATGGAGTACCGACTGGAGTCATATGCAAGCCAAAGCATTAAGTG 1860
Qy 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluLysAsnAspGlnMet 640
Db 1861 AGGAAATCCTTAATATGAGCTTGGAGATTAAGCATATGAAATGCAAGCAATGAATG 1920
Qy 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660
Db 1921 TACCTGTTCCATCTCTGTGATATGCTATGAGGCCGTCTTTTAAAGTAAATAAT 1980
Qy 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680
Db 1981 CAGATGATCTTTTGGGAGAGAGATGCGAGTGGCTAATTTGAAACCAAGATCTCC 2040
Qy 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700

```

```

Db 2041 TTTAATTTCTTGTGACCTGCACCTAAAGATGTGTGATATCATCTCTAGACCTGAGATT 2100
Qy 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2101 GAAAAGCCATCAGAGATGTCCTCCGAGGCCGTATCATATATGCTTCCGTGTGAATGACAA 2160
Qy 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2161 AGCTTAGAGTTCTTGGGATTCAGACCAACCTTGGACCTCTTACACAGCCCTGTTTC 2220
Qy 741 IleThrLeuIleValPheGlyValIleValMetGlyValIleValIleValIleLeu 760
Db 2221 ATATGCTGATGTTGTTTGGAGTTGTGATGGAGTGAATAGTGTGGCATTTGCATCTG 2280
Qy 761 IlePheThrGlyIleArgAspArgLysLysAsnLysAlaArgSerGlyLysAsnPro 780
Db 2281 AACTTCACTGGATCAGATCAGATCGAAGAAAGAAATTAAGCAAGAGTGGGAAATCT 2340
Qy 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800
Db 2341 TATGCTCATCATGATATTTAGCAAGAGAAATATATCCAGATTCCAAAACACTGATGAT 2400
Qy 801 ValGlnThrSerPhe 805
Db 2401 GTTCAGACCTCCTTT 2415

```

RESULT 3

```

LOCUS E43987 2415 bp DNA linear PAT 31-JAN-2002
DEFINITION ACE-analogous gene.
ACCESSION E43987
VERSION E43987.1 GI:18629190
KEYWORDS JP 2001046072-A/1.
SOURCE unclassified.
ORGANISM unidentified.

```

```

REFERENCE 1 (bases 1 to 2415)
AUTHORS Sugano,S. and Komatsu,T.
TITLE ACE-analogous gene
JOURNAL Patent: JP 2001046072-A 1 20-FEB-2001;
OTSUKA PHARMACEUT CO LTD

```

COMMENT

```

OS Unknown
PN JP 2001046072-A/1
PD 20-FEB-2001
PE 06-AUG-1999 JP 1999223892
PR
PI SUMIO SUGANO, TAKAMI KOMATSU
PC C12N15/09, A61K31/00, A61K31/7088, A61K38/00, A61K38/55,
A61K39/395, PC A61K39/395,
PC A61K39/395, A61K48/00, A61P9/12, C07K14/47, C07K16/08, C12N1/15, PC
C12N1/19,
PC C12N1/21, C12N5/10, C12Q1/68, G01N33/53, C12N15/00, A61K37/02, PC
A61K37/64,
PC C12N5/00
CC
FH Key 1.2415 Location/Qualifiers
FT source 1.2415
FT /organism='Unknown'.

```

FEATURES

```

source 1.2415 Location/Qualifiers
BASE COUNT 743 a 484 c 554 g 634 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 0 Length: 2415
Score: 4291.00 Matches: 805
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

```

US-09-978-385-2 (1-805) x E43987 (1-2415)

QY	1	MeerserserserftrpleuileuileusertleuValAlaValThralaAlaGlnSerThr	20
Db	1	ATGCAAGCCTTCCTGGCTCCCTTCACACTGTGGCTGACAGCTGCCTCAGCCACC	60
QY	21	IIeGIuGIuGlnAlaYsThrPheLeuAspLysPheAsnHisGluAlaGlnAspLeuPhe	40
Db	61	ATTGAGAACAGGCGCAAGACATTTTGGACAGGTTTAAACCGAAGCGGAGACCTGTTT	120
QY	41	TyrGlnserSerleuAlaSerTrpAsnTrpAsnIleThrGluGluAsnValGln	60
Db	121	TATCAAGTTCACCTGGCTTCTTGGAATTTTAAACCAATTTATTCAGAGAGATGTCCAA	180
QY	61	AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrIleuAla	80
Db	181	AACATGATATATGCTGGGAGCAAAATGTCCTGCTTTTAAAGAACACATCCACTTCC	240
QY	81	GlnMetYrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu	100
Db	241	CAATGATATCCACTACAGAAATTCAGAAATTCACAGCTCAACCTGCAGCTGCCTT	300
QY	101	GlnGlnAsnGlySerValIleuSerGlnAspLysSerLysArgLeuAsnThrIleLeu	120
Db	301	CAGCAAAATGGCTTCAGGCTCTCAGAGCAAGCAAGCAAAAGGTGTAACCAATTTCTA	360
QY	121	AsnThrMetSerThrIleYrSerThrGlyLysValCysAsnProAspAsnProGlnGlu	140
Db	361	AATACATGAGCAACCATCTACACTACTGGAAAAAGTTGTAAACCCAGATATCCACAA	420
QY	141	CysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGlu	160
Db	421	TGCTTATATCTTACACAGGTTTGAAATGAAATAATGGCAACAGTTTGACTACATAG	480
QY	161	ArgLeuTrpAlaTrpGlnSerTrpAsnSerGluValGlyLysGlnLeuArgProLeuYr	180
Db	481	AGGCTCGGGCTTGGGAAACCTGGAATCTGAGGTCGGCAACACACTGAGGCCATTAT	540
QY	181	GluGluTrpValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAspTyrGly	200
Db	541	GAAACAGATATGCTGCTGAAATAATGAGCGCAAGCAACAACTTATAGAGCACTATGG	600
QY	201	AspYrTrpAlaTrpGlyAspTyrGluValAlaAsnGlyValAspGlyTyrAspTyrSerTrgly	220
Db	601	GATTATTTGGAGAGAGACACTTGAAGTAAATGGGGTATGATGCTTATGACTACAGCGCGGC	660
QY	221	GlnLeuIleGlnAspValGlnHisThrPheGlnGlnIleLysProLeuYrGluHisLeu	240
Db	661	CAGTTGATGAAGTGGAGAACATACCTTTGAAGAGATTAAACCATTTATATGAACATCTT	720
QY	241	HisAlaIyrValArgAlaLysLeuMetAsnAlaIyrProSerTyrIleSerProIleGly	260
Db	721	CATCCCTATGTGAGGGCAAAATGTGAATATGCTTACCTTCTTATATCAGTCCAAATTTGA	780
QY	261	CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuYrSer	280
Db	781	TGCCCTCCCTGCTATTTGCTTGGTATATGTGGGGTAGATTTTGGACAAATCTGACTCT	840
QY	281	LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln	300
Db	841	TTTGACATTCCTCTTGGACAGAAACCAACATATGATGTATACATGATGATGGTGGACAG	900
QY	301	AlaIleTrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerValGlyLeu	320
Db	901	GCCCTGGGATGACGAAATATTCAAAGAGAGCGGCAAAATCTTTGTATCTGTGGTCTT	960
QY	321	ProAsnMetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGlyAsnValAln	340
Db	961	CTTATATATGACTCAAGGATCTTGGGAAATTCATGCTTACAGGACCCAGAAATGTTGAG	1020
QY	341	LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet	360
Db	1021	AAACCAATCTGCATCCCAACAGCTTGGAGCTGGGGGAAAGCGGACACTTACAGATCTCTTATG	1080

QY	361	CysThrIysValIThrMetLaspPheLeuThrAlaHisHISGLMetGlyHISrIleGln	380
Db	1081	TCGCAAAAGGTGCAATATGGACGACTTCCTGACAGCTCATCATGAGATGGGCATATCCAG	1140
QY	381	TyrAspMetAlaTyrAlaIaIaGlnProPheLeuIleuTyrAsnGlyAlaAsnGluGlyPhe	400
Db	1141	TATGATATGGCATATGCTGCAACAACCTTTTCTGCTAAGAAATGAGAGCTAATGAAGATTC	1200
QY	401	HisGluAlaValIGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer	420
Db	1201	CATGAAGCTGTGGGGAAATCATGATCACTTTCGACGCCACACCTPAACCATTTAAATCC	1260
QY	421	IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThiGluIleAsnPheLeuLeu	440
Db	1261	ATGTGCTCTCTGTACCCGATTTTTCAGAAGAACATGAAGAACGAATATACTCTCGTC	1320
QY	441	LysGlnAlaLeuThrIleValIGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg	460
Db	1321	AAACAGACACTCACGATTTGTGGAGCTCGCATTTACTTACATGTGTAGAACAGTGGAGG	1380
QY	461	TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysTrpTrpGlnMet	480
Db	1381	TGGATGGCTTTAAAGGGAAATTCCTCAAAAGCACTGATGATPAAAAAGTGGGAGATG	1440
QY	481	LysArgGluIleValIGlyValIaIaGluProValProHisAspGluThrTyrCysAspPro	500
Db	1441	AAGCAGAGATAGTGGGGGCGTGGAAACCTGTGCCCATGATGAAGAACTACTGTGACCCC	1500
QY	501	AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu	520
Db	1501	GCATCTCGTGTCCATGTGTTCTTAATGATATCTCATTCATTCGATTAATTAACAAGACCTT	1560
QY	521	TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHISGLIProLeuHis	540
Db	1561	TACCATTTCCAGTTTCAAGAAGCACTTTGTCAAGCAGCTAAACATGAAGGCCCTCTGCAC	1620
QY	541	LysCysAspIleSerAsnSerThrGluIaIaGlyGlnLysLeuPheAsnMetLeuArgLeu	560
Db	1621	AAATGTGACATCTCAAACTCTACAGAAGCTGGACACAAACTGTTCAAATATGCTGAGGCTT	1680
QY	561	GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValIaIaLysAsnMetAsn	580
Db	1681	GGAAATCATGACCCCTGGACCTTACGATTTGGAAATGTTGTGGAGCACAAGACATGAT	1740
QY	581	ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys	600
Db	1741	GTAAGGCCACTGCTCAACTACTTTGAGCCCTTATTACTGCGTGAAGACAGACAAG	1800
QY	601	AsnSerPheValIGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal	620
Db	1801	AATCTTTTGGGATGAGATACCGATGGAGTCCATAAGCAGACCAAGAATCAAAATG	1860
QY	621	ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet	640
Db	1861	AGGATAAACCTTAAATATAGCTCTTGGAGATAAAGCATATGATGGACAGACATGAATG	1920
QY	641	TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn	660
Db	1921	TACCTGTTCGATCATCTGTGTGCAATATGCTATGAGCAGCATACTTTTAAAGTAAAAAT	1980
QY	661	GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer	680
Db	1981	CAGATGATCTTTTGGGAGAGAGATGTCGAGTGGCTAAATTTAAACCAAGAAATCTCC	2040
QY	681	PheAsnPhePheValITrrAlaProLysAsnValSerAspIleIleProArgThrGluVal	700
Db	2041	TTTATATTCTTGTGCACTGCACTCAAAATATGTCGTGATATCATTCCTTAGAACGAAGTT	2100
QY	701	GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn	720
Db	2101	GAATAAGGCATACAGATGCTCCGAGCCGATCAATGAATGCTTCCGTCTCAATGACAAAC	2160

```

Qy 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2161 AGCTTACAGAGTTCTGGGGATACAGCCAACTGGACCTTCTTAACAGAGCCCTGTTCC 2220
Qy 741 IleThrLeuIleValPheGlyValIleMetGlyValIleValValGlyIleValIleLeu 760
Db 2221 ATATGGCTGATGTTGTTGGAGTTGATGGAGTATGATGGTGGCATGTGCATCTCG 2280
Qy 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780
Db 2281 ATCTTCACTGGATCAGAGATCGGAGAGAAAAATTAACAGAACTGGAGAAATCTCT 2340
Qy 781 TyrAlaSerIleAspIleSerLysGlyLysAsnProGlyPheGlnAsnThrAspAsp 800
Db 2341 TATGCCCTCCATCATATTAACAAGAGAAATAATCAAGATTCACAAACACTGATGAT 2400
Qy 801 ValGlnThrSerPhe 805
Db 2401 GTTCAGACCTCTCTT 2415

RESULT 4
LOCUS E39033 2418 bp DNA linear PAT 18-JUN-2001
DEFINITION MPROT15 polypeptide and MPROT15 polynucleotide.
ACCESSION E39033.1 GI:13017695
VERSION JP 1999318472-A/1.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2418)
AUTHORS Christopher D. S. and Nicola B.
TITLE MPROT15 polypeptide and MPROT15 polynucleotide
JOURNAL Patent: JP 1999318472-A 1 24-NOV-1999.
SMITHKLINE BEECHAM CORP PUBLIC LTD CO
COMMENT OS Homo sapiens (human)
PN JP 1999318472-A/1
PD 24-NOV-1999
PE 22-JAN-1999 JP 1999014949
PR 13-MAY-1998 GB 9810373.2, 18-AUG-1998 GB 9818009.4 PT
CHRISTOPHER D SAZAN, NICOLA BAGESU
PC C12N15/09, A61K31/00, A61K31/00, A61K31/00, A61K31/70,
PC A61K38/00,
PC A61K39/395, A61K45/00, A61K48/00, C07K14/47, C12P21/02, C12Q1/68,
PC G01N33/50, G01N33/50, G01N33/566, G01N33/68//C12P21/02,
C12R1:91, PC C12N15/00,
PC A61K37/02
CC
FH Key Location/Qualifiers
FT source 1..2418
Location/Qualifiers
/organism='Homo sapiens (human)'.
FEATURES
source 1..2418
Location/Qualifiers
/organism='Homo sapiens'
BASE COUNT 744 a 484 c 555 g 635 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2418
Score: 4291.00 Matches: 805
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6
US-09-978-385-2 (1-805) x E39033 (1-2418)
Qy 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
Db 1 ATGTCAAGCTCTTCTGGCTCTTCTCAGCCTTGTGTTGTAACGTCTGCTCAGTCCACC 60

```

```

Qy 21 IleGluGluAlaIleLysThrPheLeuAspLysPheAsnIleGluAlaGluAspLeuPhe 40
Db 61 ATTGAGGAACAGGCCAGACATTTTGGACAAGTTTAAACACAGAACGCCAGACTGTCTC 120
Qy 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsnValGln 60
Db 121 TATCAAGCTTCACTTCTCTTGGAAATTAACCAATATTAATGTAAGAAAGTCCAA 180
Qy 61 AsnMetAsnAspAlaGlyAspLysTrpSerThrPheLeuGlyGlnIleSerThrLeuAla 80
Db 181 AACATGAATTAATGCTGGGGACAAATGCTCTCTTTTAAAGAAACGTCCACACTGGCC 240
Qy 81 GlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLysGlnIleGluAlaLeu 100
Db 241 CAATGTATCCACTCAAGAAATTAATTCAGATTCACAGTCACAGCTTCAGCTCAGCTCTT 300
Qy 101 GlnGluAsnGlySerSerValLeuSerGlyLysAspLysSerLysArgLeuAsnThrIleLeu 120
Db 301 CAGCAAAATGGGCTTCTCAGTCTCTCAGAAAGCAAGCAAAACGTTGAACACAAATCTTA 360
Qy 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
Db 361 AATTAACAATGAGCACCATCTACAGTACGAAAGTTGTATACCCAGATAATCCACAAGAA 420
Qy 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
Db 421 TGCATTATTACTTGAACCAAGCTTTGATGAATTAATGCAAAACATTTAGACTACAAATGAG 480
Qy 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGlyValAlaGlyLysGlnLeuArgProLeuTyr 180
Db 481 AGGCTCTGGGGCTTGGGAAAGCTGGAGATCTGAGTGGGCAAGCAGCTGAGCCATTTAT 540
Qy 181 GlnGluTyrValValLeuLeuLysAsnGluMetAlaArgAlaAsnIleTyrGluAspTyrGly 200
Db 541 GAAGAGTATGTGCTTAAAAAATGAGATGGCAAGCAAGCAATCTTATGAGACTATGAGG 600
Qy 201 AspTyrTrpArgGlyAspTyrGlyValAlaAsnGlyValAspGlyTyrAspTyrSerArgGly 220
Db 601 GATTATTGGAGAGAGACATATGAAAGTAAATGGGGTAGATTTGGACAATCTGTACTCT 660
Qy 221 GlnLeuIleGluAspValGlnIleThrPheGlnGluIleLysProLeuTyrGlnIleLeu 240
Db 661 CAGTGTATTTGAAGATGTGGAACATATCTTTGAAGAAGATTAACCATATATGAAACATCTT 720
Qy 241 HisAlaTyrValAlaArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db 721 CATGCTATGTGAGGGCAAGATGATGATGATGATCTTATCTCTTAATCAGTCAATTGGA 780
Qy 261 CysLeuProAlaIleLysLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280
Db 781 TGGCTCCCTGCTCATTTCTCTTGGATATATGGGGGTGATTTTGGACAATCTGTACTCT 840
Qy 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 841 TTGACAGTTCCTTGGACAAACCAAAACATATAGATTTCTGATGCAAAATGGTGGACAG 900
Qy 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGlyLysPhePheValSerValGlyLeu 320
Db 901 GCCTGGATGCACAGACATATTCAGAGAGGCGAGAAAGTCTTGTATCTGTTGGCTCTT 960
Qy 321 ProAsnMetThrGlnGlyPheTrpGlyAsnSerMetLeuThrAspProGlyAsnValGln 340
Db 961 CCTATATGACTCAAGAGTTCTGGGAAATTCATGCTAACGAGCAACGAGAAATGTTCCAG 1020
Qy 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1021 AAAGAGCTCCCATCCACAGCTTGGGACCTTGGGAGAGGCGCACTTCAGAGATCTTATG 1080
Qy 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGlnMetGlyHisIleGln 380
Db 1081 TGCACAAAGGTGACAAATGAGACGACTCTCTGACAGCTCATGATGATGGGCAATTCAG 1140

```

QY 381 TyrAspMetAlaTyrAlaIaIaGlnProPheLeuLeuAArgsnGlyAlaAsnGluGlyPhe 400
 |||||
 Db 1141 TATGATATGGCATATGCTGGACAAACCTTTCTGCTAGAAATGAGCATATGAAGATTC 1200
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 |||||
 Db 1201 CATACAGCTGTGGGAAATCATCTCACTTCTGCACCCACACCTAAGCATTTAAATCC 1260
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440
 |||||
 Db 1261 ATTGCTCTTCTGTCAACCCGATTTTCAAGAGACATGATAAAGAAATTAACCTTCCTC 1320
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrIleMetLeuGluLysTyrArg 460
 |||||
 Db 1321 AAACAGACCTCAGCATTTGGGCTCTGCATTTACTTACATGTTAGAGAACTGGACG 1380
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysTyrTrpGluMet 480
 |||||
 Db 1381 TGGATGGCTTTAAAGGGGAATTTCCCAAGACAGTGGATGAAAAAGTGGTGAGATG 1440
 QY 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500
 |||||
 Db 1441 AAGCGAGAGATGTTGGGGTGTGGAACTGTGCCCATGATGAACATACCTGTGACCC 1500
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520
 |||||
 Db 1501 GCATCTCTGTCATGTTTCTTAATGATTACATTCATTCGATTTACACAGAACCTT 1560
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaAlaLysHisGluGlyProLeuHis 540
 |||||
 Db 1561 TACCAATTCAGATTCAAGAGACACTTGTCAAGCAGCTAAACATGAAGGCCCTCTGCAC 1620
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGluGlnLysLeuPheAsnMetLeuArgLeu 560
 |||||
 Db 1621 AAATGTGACATCTCAAACTGACAGAACTGTGACAACTGTCAATATCTGAGGGCTT 1680
 QY 561 GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValAlaGlyAlaLysAsnMetAsn 580
 |||||
 Db 1681 GGAAATATCAGAACCTGGACCTTACATTTGGAATGTTGAGAGCAAAAGACATGAT 1740
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpPheLysAspGlnAsnLys 600
 |||||
 Db 1741 GTAAAGCCACTGCTCACTACTTTAGGCCCTTATTTACCTGCTGAAAGACAGAACAG 1800
 QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620
 |||||
 Db 1801 AATCTTTTGGGATGGAGTACCGACGTGAGTCCATATGACGACCAAGCATCAAATG 1860
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet 640
 |||||
 Db 1861 AGGATAGCCTAAATACAGCTCTTGAGATTAAGCATATGAAATGAGACATGAATG 1920
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660
 |||||
 Db 1921 TACCTGTCCTGATCATCTGTCATATGCTATGAGCGACGACTTTTAAAGATAAAAT 1980
 QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgLysSer 680
 |||||
 Db 1981 CAGATGATCTTTTGGGAGGAGATGTCGAGTGGCTAATTTTAAACCAAGAACTTCC 2040
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700
 |||||
 Db 2041 TTTAATTTCTTGTGCTACCTACCTAAATATGTCGATATCATCTTGAAGACTGAAGT 2100
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
 |||||
 Db 2101 GAAAGAGCCATCAGATGTCGCGAGACCTATCAATGATGATCTTCGCTGATGACAAAC 2160
 QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
 |||||
 Db 2161 AGCTTGAAGTTCGTGGGATACAGCCACACCTTGACCTCTTACACAGGCCCTGTTC 2220
 QY 741 IleTrpLeuIleValPheGlyValAlaMetGlyValIleValIleValIleValIleLeu 760

Db 2221 ATATGGCTATGTTTGTGGATTTGATGGAGATAGTGTGTCATGTCATCCTG 2280
 |||||
 QY 761 IlePheThrGlyIleArgAspArgLysLysAsnLysAlaArgSerGlyGluAsnPro 780
 |||||
 Db 2281 ATCTTCACGTGGATCGAATCGAAGAGAAATTAAGCAAGAGTGGAGAAATTCCT 2340
 QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAspProGlyPheGlnAsnThrAspAsp 800
 |||||
 Db 2341 TATGCTTCATGATTTTGCAAGAGAGAAATATATCCGATTCGAAACGATGATGAT 2400
 QY 801 ValGlnThrSerPhe 805
 |||||
 Db 2401 GTTCAGACCTCCTTT 2415
 RESULT 5
 E4398
 LOCUS E4398 2599 bp DNA linear PAT 31-JAN-2002
 DEFINITION ACE-analogous gene.
 ACCESSION E4398
 VERSION 1 GI:18629191
 KEYWORDS JP 2001046072-A/2.
 SOURCE JP 2001046072-A-2.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 2599)
 AUTHORS Sugano,S. and Komatsu,T.
 TITLE ACE-analogous gene
 JOURNAL Patent: JP 2001046072-A 2 20-FEB-2001;
 OTSUKA PHARMACEUT CO LTD
 COMMENT OS Unknown
 PN JP 2001046072-A/2
 PD 20-FEB-2001
 PF 06-AUG-1999 JP 1999223892
 PR SUMIO,SUGANO,TAKAMI,KOMATSU
 PI C12N15/09,A61K31/00,A61K31/77088,A61K38/00,A61K38/55,
 PC A61K39/395,PC A61K39/395,
 PC A61K39/395,A61K48/00,A61P9/12,C07K14/47,C07K16/08,C12N1/15,PC
 C12N1/19,
 PC C12N1/21,C12N5/10,C12Q1/68,G01N33/53,C12N15/00,A61K37/02,PC
 A61K37/64,
 PC C12N5/00
 CC CC
 FH FH
 FT CDS
 FEATURES
 source 1..2599 location/Qualifiers
 /db_xref="taxon:32644"
 BASE COUNT 807 a 503 c 591 g 698 t
 ORIGIN
 Alignment Scores: 2599
 Pred. No.: 0
 Score: 4291.00
 Percent Similarity: 100.00% Matches: 805
 Best local Similarity: 100.00% Conservative: 0
 Query Match: 100.00% Mismatches: 0
 DB: 6 Indels: 0
 Gaps: 0
 US-09-978-385-2 (1-805) x E4398 (1-2599)
 QY 1 MetSerSerSerTrpLeuLeuLeuSerLeuValAlaValThrAlaIaIaGlnSerThr 20
 |||||
 Db 55 ATGTCAACCTTCTCTGCTCTCTCTCAAGCTTTCTGTAACGTGCTGCTCAAGTCCACC 114
 QY 21 IleGluGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
 |||||
 Db 115 ATTGAGAAAGGCCAAGACATTTTGGACAAAGTTTAAACACGAGAAAGCCGAGACCTGTT 174
 QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrTrpAsnThrAsnIleThrGluGluAsnValGln 60
 |||||

Db 175 TATCAAGTCTGCTGCTTGGATTAATACCAATATTAAGTGAAGATGTCAC 234
QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuYsgLugInserThrLeuAla 80
Db 235 AACATGATATATGCTGGGACCAATGCTGCTTTTAAAGAACAGTCCACATTTGCC 294
QY 81 GluMetLysTrpProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
Db 295 CAAATGATACCTACACAGAAATTCAGAAATTCACAGTCAAGCTCAGCTGACGCTCT 354
QY 101 GlnGlnAsnGlySerValLeuSerGlnAspLysSerLysArgLeuAsnThrIleLeu 120
Db 355 CAGCAAAATGGGTCTTCAGTCTCTCAGAAAGACAGACAAAGCGTTGAACACAAATTC 414
QY 121 AsnThrMetSerThrIleLeuSerThrGlnLysValLysAsnProAspAsnProGlnGlu 140
Db 415 AATACATATGACACCATCTACAGTACAGTGAAGATTGTAACCCACATATTCACAGAA 474
QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTrpAsnGlu 160
Db 475 TCGTTATTAATCTGACACAGGTTGAATGAATATGCAACAGTTTACACTACATGAG 534
QY 161 ArgLeuTrpAlaTrpLysSerTrpArgSerGluValGlyLysGlnLeuAspProLeuTrp 180
Db 535 AGGCTCTGGGCTGGGAAAGCTGGAGATCTGAGCTGGCAGACAGCTGAGCCATTAAT 594
QY 181 GlnGluTrpValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGlnAspTrpGly 200
Db 595 GAAGATATATGCTGCTTGAAGAAATGAGATGGCAGACGAAATCATATGAGGACATATGG 654
QY 201 AspTyrTrpArgGlyAspTrpGluValAsnGlyValAspGlyTyrAspTrpSerArgGly 220
Db 655 GATATATGAGAGAGACATATGAAATGAATGGGATGAGTGGCTATGACACAGCCGCGC 714
QY 221 GlnLeuIleGlnAspValGlnHisThrPheGlnGluIleLysProLeuTyrGlnHisLeu 240
Db 715 CAGTGAATGAGATGTCGACATACCTTTGAAGATTAACCTTTATATGACATCTT 774
QY 241 HisAlaTrpValAlaAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db 775 CATGCCATATGAGGACAAAGTTGATGATGCTTCCTTATATACAGTCCAAATTTGA 834
QY 261 CysLeuProAlaHisLeuLeuGlnLysMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280
Db 835 TCCCTCCCTGCTCATTTGCTGCTGATATGAGGCTAATTTTGACAAATCTGACTCT 894
QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 895 TTGACAGTTCCCTTTGGACAGAAACAAACATAGATTTACTGATGCAATGCTGAGACG 954
QY 301 AlaTrpAspAlaGlnArgIlePheLysGlnAlaGluLysPhePheValSerValGlyLeu 320
Db 955 GCGTGGATGACAGAGAAATATTCAGAGAGCCGAGAAAGTCTTTGATCTGGTTGGTCTT 1014
QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
Db 1015 CCTATATAGCTCAAGGATCTCGGAAATTCATGCTAACGAGCCAGGAAATGTTCCAG 1074
QY 341 LysAlaValLysHisProThrAlaTrpAspLeuGlnLysGlyAspPheArgIleLeuMet 360
Db 1075 AAAGCAGTCTGCCATCCACAGCTTTGGAGACCTGGGGAAGGCGACTTCAGATCTTAATG 1134
QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGlnGluMetGlyHisIleGln 380
Db 1135 TGCAACAAGGTGACAAATGAGAGACTCTCTGACAGCTCATATGATGAGGGGCAATTCAG 1194
QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPhe 400
Db 1195 TATATATATGCGATATGCTGCAACACTTTTCTGCTAAGAAATGAGAGCTATATGAGATT 1254
QY 401 HisGlnAlaValGlnGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
Db 1255 CATACAGCTGTGGGAAATCATGCTCTTTCGACGCCACACCTTAAGCATTTAAATCC 1314
QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440
Db 1315 AATGGCTCTTCTGACCCGATTTTTCAGAGACATATGAAGAAATTAACCTCTGCTC 1374
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheTrpThrMetLeuGlnLysTrpArg 460
Db 1375 AATCAAGACATCAGATTTGTTGGAGCTGCGCATTTACTTACATGTTAGAAATGAGAG 1434
QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrpGlnMet 480
Db 1435 TGGATGCTCTTAAAGGGAAATTCACAAAGCAGTGGATGAAAGAAAGTGGGAGATG 1494
QY 481 LysArgGluIleValGlyValValGluProValProHisArgGluThrTyrAspPro 500
Db 1495 AAGCAGAGATATGGGGGTGGTGGAGACCTTGGCCCATGATGAACATCTGAGACCC 1554
QY 501 AlaSerLeuPheHisValSerAsnAspTrpSerPheIleArgTyrTrpArgThrLeu 520
Db 1555 GCATCTGCTTCATGCTTTCTATGATGATTCATTCATTCGATTTTACACAGAGACCTT 1614
QY 521 TyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGlnGluProLeuHis 540
Db 1615 TACCAATTCAGTTTCAAGAAAGCACTTTGTCACAGCAGCTAAACATGAAGCCCTGAC 1674
QY 541 LysCysAspIleSerAsnSerThrGlnAlaGlyLysLeuPheAsnMetLeuGluLeu 560
Db 1675 AATATGACATCTCAAACTCTACAGAAAGCTGGACAGAAACTGTTCAATATGCTGAGGCTT 1734
QY 561 GlyLysSerGluProTrpThrLeuAlaLeuGlnAsnValValGlyAlaLysAsnMetAsn 580
Db 1735 GGAATAATCAGAAACCTTGACCCCTGATGGAATATGTTTATGAGACAAAGAACATGAA 1794
QY 581 ValArgProLeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuLysAspGlnAsnLys 600
Db 1795 GTAAGGCCATGCTCAACTCTTTGACCCCTTTTACCGCTGAAAGCCAGACAAAG 1854
QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620
Db 1855 AATCTTTTGGAGATGAGATACCGAGCTGAGCTCATATGACACAAAGACATCAAGT 1914
QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluThrAsnAspAsnGluMet 640
Db 1915 AGGATTAAGCCTTAATACAGCTCTTGGAGATTAAGCATATGAATGAACACATGAAT 1974
QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnThrPheLeuLysValLysAsn 660
Db 1975 TACCTGTCCGATCATCTGTTCATATGCTATGAGGACAGTACTTTTAAAGTAAATAAT 2034
QY 661 GluMetIleLeuPheGlnGlnGluAspValArgValAlaAsnLeuLysProArgIleSer 680
Db 2035 CAGATATCTCTTTGGGAGGAGATGTCGAGTGGCTGAATTTGAACCAAGATCTCC 2094
QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleLeuProArgThrGluVal 700
Db 2095 TTTAATTTCTTCTGCTACCTCACCCTTAATAATGCTGATATCATCTTAAGACTAAGTT 2154
QY 701 GlnLysAlaIleArgMetSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2155 GAAAGGCGCATGAGATGTCGCGGAGCCCTATCATGATGCTTTCCGCTGTGATGACAC 2214
QY 721 SerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2215 AGCTTACAGTCTTCTGGGATACAGCAACACTTGGACCTCTTAACACACCCCTGTTCC 2274
QY 741 IleThrLeuIleValPheGlyValValMetGlyValIleValValGlyIleValIleLeu 760
Db 2275 AATATGCTGATGTTTGTGAGATTTGATGAGGAGATAGGCTGATGCTGATCTGCTG 2334
QY 761 IlePheThrGlyIleArgAspArgLysLysAsnLysAlaArgSerGlyGluAspPro 780
Db 2335 ATCTTCACAGGATCAGAGATGCGAAGAAAGAAATTAAGCAAGAGTGGAGAAATCT 2394

QY 781 Tyra1aSer1leaspl1eSer1ySg1yLu1aSnasProG1yPheG1naSnThrAsp800
 Db 2395 TATGCTTCATGATATATAGAAAGAAATATATCCAGATTCCAAAACTGATGAT 2454
 QY 801 ValGlnThrSerPhe 805
 Db 2455 GTTCAGACTCTCTTT 2469

RESULT 6
 AF291820 3325 bp mRNA linear PRI 15-JUN-2001
 LOCUS Homo sapiens ACE-related carboxypeptidase ACE2 mRNA, complete cds.
 DEFINITION AF291820
 ACCESSION AF291820.1 GI:9802432
 VERSION
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 3325)
 Donoghue, M., Hsieh, F., Baronas, E., Godbout, K., Gosselin, M.,
 Stagliano, N., Donovan, M., Woolf, B., Robison, K., Jeyaseelan, R.,
 Breitbart, R.E. and Acton, S.
 A novel angiotensin-converting enzyme-related carboxypeptidase
 (ACE2) converts angiotensin I to angiotensin 1-9
 Circ. Res. 87 (5), E1-E9 (2000)

JOURNAL MEDLINE
 20429895
 PUBMED 10969042
 REFERENCE 2 (bases 1 to 3325)
 Donoghue, M., Woolf, B., Robison, K. and Acton, S.
 Direct Submision
 Submitted (01-AUG-2000) Cardiovascular Biology, Millennium
 Pharmaceuticals, Inc, 75 Sidney Street, Cambridge, MA 02139, USA

FEATURES
 source location/Qualifiers
 1..3325
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 21..2438
 /codon_start=1
 /product="ACE-related carboxypeptidase ACE2"
 /protein_id="AAF9721.1"
 /db_xref="GI:9802432"
 /translation="MSSSSWLLSLVAVTAQSTIEEQAKTFIDKFNHEADLFYQSS
 LASVNTNITEENYONMNMAGDKMSAFKEQSTIAQWLPQIETONLTVKIQIALQO
 NGSSVLEDESKRINTILNTNSTIYSGTKGCPDMPQELILEPLNTEMANSLDYNE
 RLAMNESMSPEYKQRLREYEVYVLEKEMARANTIEDIGDWRGDYEVNGVDYDS
 RQGLIEDVEHTEETKPYEHLHAYVRAKLMAVPSYISPIGCLPAHLGEMWGRFT
 NLYSLVPEPGQPNIDVDAMVDAQDFKEKEKFEFVSGLPMTQGEWGRFT
 DPGAVQKACHPETAMDLDKGFRIIMCTKRYDDDLTAHHEMGILOYMAYAAOPEFL
 RANGANGFEHVAIGEINLSAATPKHLKSTGLSPQFODNETEINFLIKOALTIVGTL
 PPTYLEKRWKMFKEGTEIPKQDMKMKWEMKREYGVPEVPHDTCYDPPASLEFVSN
 DSYFIRYRTITQPOFOBALCOAKKHGEPFLKCDISTEAGOKLFNMLRKGSEPW
 TTALENVGAKKMKNVRLNTEFPLETWLKQDNKNSFGWSDSPVADQSLKAVISL
 KSLADKAIEMWDMENMYLERSSVAYAMQVLELVKNOMILGEEEDVRYANLKPISFN
 FVTAPKNSVDIIPTEVEKAIKMSRSRINDAFRLNDMSLELIGIOTPLGPPNPPVS
 IMLIEGVGVGIYVGIYVILFTGIRDKKKKKKANSNPVASTDISGNNPGRONT
 DDVQTSF"

BASE COUNT 1022 a 637 c 747 g 919 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3325
 Score: 4291.00 Matches: 805
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-978-385-2 (1-805) x AF291820 (1-3325)

QY 1 MetSerSerSerThrPheLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 21 ANGTCAAGCTCTTCTGCGCTCTTCACGCTTGTGCTGTAAGTCTGCTACGCC 80
 QY 21 TleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 81 ATTGAGAACAGAGCCCAAGACATTTTGGACAGATTAAACAGAGAACCCGAGACCTGTTC 140
 QY 41 TyrGlnSerSerLeuAlaSerThrPasnTyrAsnThrAsnTleThGluGluAsnValGln 60
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 141 TATCAAGTCTACCTGCTCTTGGAAATTAATACCAATATTTACTGAAGAGAAATGCCAA 200
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 201 AACATGAATAAATGCTGGGACAAATGCTGCTTTTAAAGAACATGCCACTTCCC 260
 QY 81 GlnMetTyrProLeuGlnGluTleGlnAsnLeuThrValLysLeuGlnAlaLeu 100
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 261 CAATATGATCCACTACACAAATAATCAGATCTCACAGTCAAGCTTCAGCTGAGGCTCTT 320
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerArgLeuAsnThrTleLeu 120
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 321 CAGCAAAATGGGTCTTCACTGCTCTCAGAGACAAAGACAAAGGTTGAACCAATTTCTTA 380
 QY 121 AsnThrMetSerThrTleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 381 AATACAAATAGACACCATCATGACTGGAAGAAAGTTTGAACCCGATTAATCCACAAAGAA 440
 QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGluTleMetAlaAsnSerLeuAspTyrAsnGlu 160
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 441 TCGTTATTAATCTTGAACCAAGGTTTGAATGAATAATAGGCAACAGTTTGAATACATAGAG 500
 QY 161 ArgLeuThrPalaThrGlnSerThrParSergLysValGlyLysGlnLeuArgProLeuTyr 180
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 501 AAGCTCTGGGCTTGGGAAAGCTGAGATCTGAGGTCGACAGACACTGAGCCATTAAT 560
 QY 181 GluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 561 GAAGAGTATGCGTCTTGAATAATGAGATGGCAAGACGAATCATTAAGAGACTATGCG 620
 QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 621 GATTATTGGAGAGACTATGAAATGAATGGGAGATGAGTATGATACAGCCGCGCG 680
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluTleLysProLeuTyrGlnHisLeu 240
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 681 CAGTTGATTGAAGATGTGGAACATACCTTTGAAGAGATTAAACCAATTATGAACATCTT 740
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 741 CATGCCATGTGAGGCGCAAGTGAATGATGCTATCTCTTATATCAGTCCAAATTCGA 800
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpAsnLeuTyrSer 280
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 801 TGCCCTCCCTGCCATTTCTTGGTATGTGGGGTAGATTTTGGACAAATCTGTACTCT 860
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 861 TTGACAGTTCCTTTTGGCAGAAACCAACATAGTGTACTGTAGTGCATGTGGACAG 920
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 921 GCCTGGATGCACAGAGAAATATTCAGAGAGGCCGGAAGTCTTGTATATCGTTGGCTTT 980
 QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 981 CCTAATATGACTCAAGGATTTCTGGGAAATTCATGCTTAACGGACCCACGAAATGTTTCAG 1040
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1041 AAAGCAGTCTGCAATCCACAGCTTGGGACCTGGGGAAGGGCGACTTCAGATCTTATG 1100
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1101 TGCAAAAGGTGACAAATGAGACACTCTCTGACACTCATGTGAGATGGGCAATATCCAG 1160

QY	21	IIleGIuGIuAlaIaIyStrPheLeuAspLySPheAsnHisGIuAlaGIuAspLeuPhe	40
Db	95	ATTGAGAACAGGCCAAGACATTTTGGACAAGTTTAACCGAAGGCCAAGACCTGTTTC	154
QY	41	TyrGIaSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIlePheGIuGluAsnValGIu	60
Db	155	TATCAAGTTCACCTTCCTTGGAATTTAAACCCAAATTTACGAAAGATGTCCAA	214
QY	61	AsnMetAsnAsnAlaGIuAspLyStrPserAlaPheLeuLyGIuGlnSerThrLeuAla	80
Db	215	AACATCAATAAGCTGGGACAATGTCTGCCTTTTAAAGAACAGTCCACACTTGCC	274
QY	81	GlnMetTyrProLeuGlnGIuIleGIuAsnLeuThrValLyLeuGIuLeuGlnAlaLeu	100
Db	275	CAAATGTATCCACTACAAAGAAATTCAGAAATCCAGATCAAGCTTCAGCTGAGCTCT	334
QY	101	GIuGlnAsnGIuSerSerValLeuSerGIuAspLySerLyArgLeuAsnThrIleLeu	120
Db	335	CAGCAAAATGGGCTTCAGTGCCTCAGAAAGAACAGCAAGCGTTGACCAATTTCTA	394
QY	121	AspThrMetSerThrIleIleTyrSerThrGIuLyValCysAsnProAspAsnProGlnIu	140
Db	395	AATACATAGAGACACCTACAGTACTGAGAAAAGTTGTAAACCCAGATAATCCAAAGAA	454
QY	141	CysLeuLeuLeuGIuProGIuLeuAsnGIuIleMetAlaAsnSerLeuAspTyrAsnIu	160
Db	455	TGCTTATTTACTGAAACACAGTTTGAAATGAATATGCAACAGCTTTAGACTACATAG	514
QY	161	ArgLeuTrpAlaTrpGIuSerTrpArgSerGIuValGIuLyGIuLeuArgProLeuTyr	180
Db	515	AGGCTCTGGGCTGGGAAAGCTGGAGATCTGAGTGGGCAAGCAGCTGAGGCCATTATAT	574
QY	181	GIuGIuTyrValValLeuLeuLysAsnGIuMetAlaArgAlaAsnHisTyrGIuAspTyrGIu	200
Db	575	GAAAGATAGTGTCTTGAAATATAGATGGCAGACGAAATCATTTAGAGAGACTAGGG	634
QY	201	AspTyrTrpArgGIuAspTyrGIuValAsnGIuValAspGIuTyrAspTyrSerArgGIu	220
Db	635	GATTTATGGAGAGGAGACTATGAACTAAATGGGCTAGATGCTATGACTACAGCGCGGC	694
QY	221	GlnLeuIleGIuAspValGIuHisThrPheGIuGIuIleLysProLeuTyrGIuHisLeu	240
Db	695	CAGTTGATTGAAGATGTGACACATCTTTGAGAGATTTAAACCATTTATGAACATCTT	754
QY	241	HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGIu	260
Db	755	CATGCTTATGGAGGGCAAGTGTGATGAATGCCCTTCTCTTATATCAGTCCAAATGGA	814
QY	261	CysLeuProAlaHisLeuLeuGIuLysPheTrpArgTyrPheTrpHisAsnLeuTyrSer	280
Db	815	TGCCCTCCCTGCTCATTTCTGTGGTATATGTGGGTAGATTTTGGACAAATCTGTACTCT	874
QY	281	LeuThrValProPheGIuLysLysProAsnIleAspValThrAspAlaMetValAspGIu	300
Db	875	TTGACAGTTCCTTGGACAGAAACCAACATAGATGTACTGATGCAATGTGTGACACG	934
QY	301	AlaTrpAspAlaGlnArgIlePheLysGIuAlaGIuLysPhePheValSerValGIuLeu	320
Db	935	GCGTGGATGTCACAGAGATATTCAGAGAGGCCGAGAACTTGTGATCTGTGGGCTT	994
QY	321	ProAsnMetTrpGlnGIuPheTrpGIuAsnSerMetLeuThrAspProGIuAsnValGIu	340
Db	995	CCTATATAGACTCAAGGATCTCTGGAAATTCATGCTAAGCGAACCCAGAGAAATGTTCA	1054
QY	341	LysAlaValLysHisProThrAlaTrpAspLeuGIuLysGIuAspPheArgIleLeuMet	360
Db	1055	AAAGCAGTCTGCGCATCCACAGCTTGGACCTGGGGAAGGGCGCTTCAGAGATCTCTTATG	1114
QY	361	CysThrLysValThrMetAspAspPheLeuThrAlaHisIleGIuMetGIuHisIleGIu	380
Db	1115	TGCACAAAGGTGACAAATGAGCAGCTTCCTGACAGCTCATCATGAGATGGGGCAATATCCAG	1174
QY	381	TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGIuAlaAsnGIuGIuPhe	400
Db	1175	TATGATATGGCATATGCTGGACAACACTTTCTGCTAGAAATGGAGCTATGAGATTC	123
QY	401	HisGIuAlaValGIuGIuIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer	420
Db	1235	CATGAACCTGTGGGAAATCATGTCACTTTCTGCAAGCCACACTTAAGCATTTAAATATCC	129
QY	421	IleGIuLeuLeuSerProAspPheGlnGIuAspAsnGIuThrGIuIleAsnPheLeuLeu	440
Db	1295	ATTGCTTCTGTCACCCGATTTTCAAGAAACAAATGAAACAAATTAACCTTCTGCTC	135
QY	441	LysGIuAlaLeuThrIleValGIuTyrLeuProPheThrTyrMetLeuGIuLysTyrArg	460
Db	1355	AAACAGACCTACAGATTTGGGAGCTCCACTTTACTTACATGTTAGAGAGTGGAGG	141
QY	461	TrpMetValPheLysGIuIleProLysAspGIuTrpMetLysLysTrpTrpGIuMet	480
Db	1415	TGGATGCTCTTTAAAGGGGAAATTTCCAAAGACAGGTGAGTGAAGAAAGTGGGAGATG	147
QY	481	LysArgGIuIleValGIuValGIuProValProHisAspGIuThrTyrCysAspPro	500
Db	1475	AAGCGAGATATGTTGGGCGTGGTGAACCTGTGCCCATGATGAACATATGCTGACCC	153
QY	501	AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu	520
Db	1535	GCATCTGTGTCATGTTTCTTAATGATTACTCATTCATTCGATATTCACAAAGGACCCCT	1594
QY	521	TyrGIuPheGlnPheGlnGIuAlaLeuCysGIuAlaAlaLysHisGIuGIuProLeuHis	540
Db	1595	TACCAATTCAGTTTCAAGAGACACTTTGCAAGCAGCTTAACATTAAGCCCTTGCCAC	165
QY	541	LysCysAspIleSerAsnSerThrGIuAlaGIuLysLeuPheAsnMetLeuArgLeu	560
Db	1655	AAATGTGACATCTCAAACTCTACAGAAAGCTGGACAGCAAACTGTCTAAATGTGTGAGGCT	1714
QY	561	GIuLysSerGIuProTrpTrpThrLeuAlaLeuGIuAsnValValGIuAlaLysAsnMetAsn	580
Db	1715	GGAATAATCAAAACCTCGAGACCCTAGCATTTGGAAAAATGTTGTAGAGCAAAAGATGAA	1774
QY	581	ValArgProLeuLeuAsnTyrPheGIuProLeuPheThrTyrLeuLysAspGIuAsnLys	600
Db	1775	GTAAGGCCATGCTCAACTACTTTGAGCCCTTATTTACTGCTGAAGACCAAGAAACAG	1834
QY	601	AsnSerPheValGIuTyrPserThrAspTrpSerProTyrAlaAspGlnSerIleLysVal	620
Db	1835	AATTTCTTTGTGGAGTGGAGTACCGACTGAGTCCATATGCAACCAAGCTCAAGTG	1894
QY	621	ArgIleSerLeuLysSerAlaLeuGIuAspLysAlaTyrGIuTrpAsnAspAsnIleMet	640
Db	1895	AGGATAGAGCTTAATAATCACCTCTTGAGAGATTAAGCATATGATGAGAACACAAATGAATG	1954
QY	641	TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLysValLysAsn	660
Db	1955	TACCTGTTCGATCATCTGTTGCATATGCTAATGAGGCACTACTTTTAAAGTAAATAAT	2014
QY	661	GlnMetIleLeuPheGIuGIuLysAspValArgValAlaAsnLeuLysProArgIleSer	680
Db	2015	CAGATATTTCTTTTGGGAGAGAGATGTGCGAAGTGGCTAATTTGAACCAAGAAATCTCC	2074
QY	681	PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGIuVal	700
Db	2075	TTTAAATTTCTTGTCTCCTCAGCTCAAAAATGTGTCTGATATATCTCTAGAACTGAAGCTT	2134
QY	701	GIuLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn	720
Db	2135	GAAGAGCCCATGAGATGTCCCGGAGCGGTATCAATGATGCTTCCGTGAATGACAAAC	2194
QY	721	SerLeuGIuPheLeuGIuIleGIuProThrLeuGIuLysProArgAsnGIuProProValSer	740
Db	2195	AGCCTAGAGTTTCTGGGAGTACAGCAACACTTGGACCTCTCAACCAAGCCCGCTGTTCC	2254
QY	741	IleTrpLeuIleValPheGIuValValaMetGIuValIleValValGIuIleValIleLeu	760

Db 2255 ATATGGCTGATTGTTTGGAGTGTGATGGAGTGTGCTGTCATCTCG 2314
 Qy 761 IIEPETHGLYILEARGSPARGLYLSYLASNLVALARGSERGLYLUNSPRO 780
 Db 2315 ATCTTACAGGATGAGATGAGAAAGAAATAAGCAAGAGTGAATAATCTT 2374
 Qy 781 TYRALSERILESPILSERLYSGLYLUAASNPROGLYPHEGLINASNTHRASPPAS 800
 Db 2375 TATGCTTCATGATATTGACAAAGAGAAATAATCCAGATTCCTCAATCATGAT 2434
 Qy 801 VALGLNTHSERPHE 805
 Db 2435 GTTCAGACCTCTT 2449
 RESULT 8
 ARI35177
 LOCUS ARI35177 3396 bp DNA linear PAT 16-MAY-2001
 DEFINITION Sequence 1 from patent US 6194556.
 ACCESSION ARI35177
 VERSION ARI35177.1 GI:14124082
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 3396)
 AUTHORS Action, S. Laurene, and Robison, K. Earl.
 TITLE Angiotensin converting enzyme homolog and therapeutic and
 diagnostic uses therefor
 JOURNAL Patent: US 6194556-A 1 27-FEB-2001;
 FEATURES
 source Location/Qualifiers
 1..3396
 /organism="unknown"
 BASE COUNT 1034 a 659 c 772 g 931 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3396
 Score: 4291.00 Matches: 805
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 DB: 6
 US-09-978-385-2 (1-805) x ARI35177 (1-3396)
 Qy 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaIaInserThr 20
 Db 82 ATGTCAAGCTCTTCTGCTCTTCTGACCTGTGCTGTAAGTCTGCTCAGTCCACC 141
 Qy 21 ILEGUGLUGLINALYSTHRPHELEUASPLYPHASNHSGLUNLAGLUNSPLEUPE 40
 Db 142 ATTGAGGAGACAGCCAAAGACATTTTGGACAACTTAAACAGACAGCCGAGCTGTC 201
 Qy 41 TYRGLINSErLeuAlaSerTrpAsnTYRAsnThrAsnIleThrGLUNASnValGln 60
 Db 202 TATCAAGTTCACCTCTCTTCTTGAATTTATACCAATATTTCTGAAAGATGTCCAA 261
 Qy 61 ASMETASnAsnIaGlyAspLysTrpSerAlaPheLeuLysGLUNInserThrLeuAla 80
 Db 262 AACCTGATATATGCTGGGACCAATGCTGCTTTTAAAGAACAGTCCACACTTGGC 321
 Qy 81 GlnMetYrProLeuGlnIleGlnAsnIleThrValLysLeuGlnInleuGlnAlaLeu 100
 Db 322 CAATATGATCTACTACAGAAATTCAGAAATTCACAGTCAAGCTCAGGCTGCTT 381
 Qy 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
 Db 382 CAGCAAAATAGTGCTTCTAGTGTGTGAGACAGCAAGACAAAGGTTGAACACAATTCYA 441
 Qy 121 AsnThrMetSerThrIleTYRserThrGlyValCysAsnProAspAsnProGlnIle 140
 Db 442 AATACAAATGAGCACCATCTACAGTACTGGAAAGTTTGTATACCAAGATTAATCCAGAA 501

Qy 141 CysLeuLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTYRAsnGlu 160
 Db 502 TGCTTATTACTTGAACACAGGTTTGAATGAATATAATGCAAAACGTTTAGACTCAATGAG 561
 Qy 161 ARGLEUTRALATPGIUSERTRTPARGSERGLUVALGLYLSGILNEUATGPROLEUTYR 180
 Db 562 AGGCTCTGGGCTTGGGAAGCTGGAGATCTGAGTCTGGCAAGACAGCTGAGCCATTATAT 621
 Qy 181 GlnGluTYRValAlaIleLysAsnGluMetAlaArgAlaAsnIleTYRGLUNAspTYRGLY 200
 Db 622 GAGAGATATGTGCTTGAATAATGACATGAGATGCGAAGACCAATCTTTAGGACTATGG 681
 Qy 201 ASPTYRTPARGIAspTYRGLUVALAsnGlyValAspGLYTYRAspTYRserArgGLY 220
 Db 682 GATTATTGAGAGAGAGACTATGAGTAAATGGGTAGATGGCTATGACTACACCCGGGC 741
 Qy 221 GlnIleIleGluAspValGlnHisThrPheGlnGluIleLysProLeuTYRGLINHisLeu 240
 Db 742 CAGTTGATTTGAAGATGTGGAAACATACCTTGAAGAGATTAACCATTAATATGAACATCTT 801
 Qy 241 HisAlaTYRValArgAlaLysLeuMetAsnAlaTYRProSerTYRILESerProIleGly 260
 Db 802 CATGCTTATGTGAGGGCAAGTTGATGAATGCCATCTCTTCTATATCAGTCCAAATTGGA 861
 Qy 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGLYARGPHETrpThrAsnLeuTYRser 280
 Db 862 TGCCCTCCCTGCTCATTTCTGTTGATATGAGGGATTTGGCAAAATCTGTACTCT 921
 Qy 281 LeuThrValLProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
 Db 922 TTGACAGTTCCTTTGGACAAACCAACCAATCAATGTCTTGTGATGCAATAGGTGGACAG 981
 Qy 301 AlaTrpAspAlaGlnArgIlePheLysGluValLysPhePheValSerValGlyLeu 320
 Db 982 GCCTGGATGACAGAGAGATTTCAAGSAGGCCAGAAAGTTCTTGTATCTGTTGGCTT 1041
 Qy 321 ProAsnMetThrGlnGlyPheTrpGLUNSerMetLeuThrAspProGlyAsnValGln 340
 Db 1042 CCTATATGACTCAAGGATTCCTGGGAAATTCATGTCTAAAGGACCCAGCAAAATGTTTCAG 1101
 Qy 341 LysAlaValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
 Db 1102 AAAGCAGTCTGCATCCACAGCTTGGGACCTGGGAGAGGCGACTTCAGAGATCTTATG 1161
 Qy 361 CysThrLysValThrMetAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380
 Db 1162 TGCAAAAGAGTACAGATGAGAGACTCTCTGACACTCATATGATGAGGGCATATCCAG 1221
 Qy 381 TYRAspMetAlaTYRAlaIleGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
 Db 1222 TATGATATGCAATATGCTGCACAAACCTTTCTGTAGAAATGAGCTAATGAAGGATTC 1281
 Qy 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaIleTrpPOLYSHISLeuLysSer 420
 Db 1282 CATGAGAGCTGTGGGGAATATGATGCTTCTGACAGCCACACCTTAAGCATTTTAAATTC 1341
 Qy 421 IleglyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440
 Db 1342 ATTGCTTCTTGCTACCCGATTTTCAAGAAGACATGAACAGAAATAAATCTCTGCTC 1401
 Qy 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTYRMetLeuGluLysTrpArg 460
 Db 1402 AAACAGACACTCACATTTGTTGGGCTCTGCCATTTTACTTTCATGTTAGAAATGTGAGG 1461
 Qy 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrpGlnMet 480
 Db 1462 TGGATGGCTTTTAAAGGGGAATTTCCAAAGACAGTGAATGAAAGTGGTGGAGATG 1521
 Qy 481 LysArgGluIleValGlyValGluLubProValProHisAspGluThrTYRCysAspPro 500
 Db 1522 AAGCAGAGATATGTTGGGTGGTGGAACTGCGCCCATGATGAACATACTGTGACCCC 1581
 Qy 501 AlaSerLeuPheHisValSerAsnAspTYRserPheIleArgTYRThrArgThrLeu 520

Db	502	TGCTATTACTGTGACACAGGTTTGCAATGAATAATGCGCAAAAGATTGAGCTACAAATGAG	561
Qy	161	ArgLeuThrAlaIrrpGluSerThrPargSerGluValGlyLysGlnLeuArgProLeuIrr	180
Db	562	AGGCTCTGGGCGCTGGGGAACCTGGAAATCTGGAGCTCGGACACACCTGAGGCCATTATAT	621
Qy	181	GluGluIrrValValLeuLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluIrrAspTyrGlu	200
Db	622	GAAAGATGATGTGGCTCTTGAAAAATGAGAGTGGCAAGACCAATCATTAATGAGACTATGGG	681
Qy	201	AspTyrThrPargLysAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgLys	220
Db	682	GATTATTGGAGAGACAGCATGAAAGTAAATGGGGTGAATGCGTATGACTACAGCGCGAGC	741
Qy	221	GlnLeuLeuGluLyspValGlnHisThrPheGlnGluIrrLeuLysProLeuTyrGlnHisLeu	240
Db	742	CAGTGTGATTGAAGATGGGGAACATACCTTTGGAAGATTAACCAATTAATGAACATCTT	801
Qy	241	HisAlaIrrValArgAlaLysLeuMetAsnAlaTyrProSerTyrLLeSerProIleuLys	260
Db	802	CATGCTCATGTGAGGCGCAAGTTGATGAAATGCCATTCCTCATATACAGTGCATTTGGA	861
Qy	261	CysLeuProAlaHisLeuLeuGlyAspMetTrrpGlyArgPheTrpHisLeuTyrSer	280
Db	862	TGCCTCCCTGCTCAATTTGCTTGGTGTGATGTGGGTAGATTTTGGACAATCTGTACCT	921
Qy	281	LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln	300
Db	922	TTTGACAGTCCCTTTGGACAGAAACCAACATGATGATTACTGATGCAATGGTGGACAG	981
Qy	301	AlaTrrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu	320
Db	982	GCGCGGATGCACAGAAATATTCAGAGGCGCGAAGAACTTTGTGTATCTGTGGTCTT	1041
Qy	321	ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln	340
Db	1042	CTTATATGACTCAAGATTTCTGGGAATTTCCATGCTTAACGCGACCGAGAAATTTGAG	1101
Qy	341	LysAlaValCysHisProThrAlaTrrpAspLeuGlyLysGlyAspPheArgIleLeuMet	360
Db	1102	AAACACAGTCTCCATCCCAACAGCTGTGGAGCTGGGGAAGCGGACCTCGAGATCCTATG	1161
Qy	361	CysThrLysValTrrpMetAspPheLeuThrAlaHisHisGlnMetGlyHisIleGln	380
Db	1162	TGCACAAAGGACAAATGACAGCTTCTGTACAGCTCATCATGAGTGGGGCATATCCAG	1221
Qy	381	TyrAspMetAlaTrrpAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluLysPhe	400
Db	1222	TATGATATGGCATATGTGTCGACAACCTTTTCTCTAAAGAAATGAGAGCTATGAAGATTTC	1281
Qy	401	HisGluAlaValGlyGluIrrMetSerLeuSerAlaAlaThrProLysHisLeuLysSer	420
Db	1282	CATGAAGCTGTGGGGAATCATGTCACTTTCGACCCACACCTGACGATTATAAATCC	1341
Qy	421	IleGlyLeuLeuSerProAspPheGlnGluLysAsnGluThrGluIrrLeuPheLeuLeu	440
Db	1342	ATTGGTCTTCTGTACCCGATTTTCAAGAAGACAATGAAGAAACGAATAAATCTCTGCTC	1401
Qy	441	LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTrrpArg	460
Db	1402	AAACACAGCACACGATTTGTGGAGCTGTGCCATTTTACTTACATGTATTAGACAGTGGAGG	1461
Qy	461	TrrpMetValPheLysGlyGluIrrProLysAsnGlnTrrpMetLysLysTrrpIrrLeuMet	480
Db	1462	TGGATGTCTTTAAAGGGGAATTTCCAAAGACCAATGATGATAAAAAAGTGGTGGAGATG	1521
Qy	481	LysArgGluIrrLeuValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro	500
Db	1522	AAGCAGAGATAGTTGGGGGTGGTGAACCTGTCCCAATGATGAAGAAACATACGTACCCC	1581
Qy	501	AlaSerLeuPheHisValSerAsnAspTyrSerPheIlaArgTyrTyrTrpThrArgThrLeu	520
Db	1582	GCATCTGTGTCACATGTTTCTAATGATTTACTCATTAATTCGATATTACACAAAGACCTT	1641

QY	521	TyrinPhegInPhegInGluAlaLeuGlySGLnAlaAlaLysSHLsgInGlyProLeuHIS	540
Db	1642	TACCAATTCAGATTTCAGAAAGACATTGTGTCAAGAGCTAAACATGAAGGCCCTCTGCAC	1701
QY	541	LysCysAspIIEserAsnSerThrGluAlaGlyInLysLeuPheAsnMetLeuArgLeu	560
Db	1702	AAATGTGCATCTCAAACTCTACAGAACTGGACGAAGAACTGTTCAATATGCTGAGGCTT	1761
QY	561	GlyIysSerGluProTrpThrLeuAlaLeuGluAsnValValAGlyAlaLysAsnMetAsn	580
Db	1762	GGAAATACAGAAACCTGGAGCCCTGACATTGGAAAAATGTTGTAGAGGCAAAACATGATGAT	1821
QY	581	ValArgProLeuLeuAsnThrPheGluProLeuPheTrpTrpLeuLysAspGlnAsnLys	600
Db	1822	GTAAGGCCACAGCTCACTAACTACTTGTAGCCCTTATTTACTTGCTGGCTGAAGACCAAGACAG	1881
QY	601	AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIIElysVal	620
Db	1882	AATTCCTTTGTGGGATGAGATGACGACACCTGGATGTCATATGCAACCAAGCATCAAACTG	1941
QY	621	ArgIIEserLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet	640
Db	1942	AGGATAACCCATAAATCATCGCTCTTGAGATTAAGCATGTGATGGAAACAGCAATGAAATG	2001
QY	641	TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn	660
Db	2002	TACCTGTTCGATCATCTGTTGTCATATCTTGAAGCGTACTTTTAAAAAGTAAAAAT	2061
QY	661	GlnMetIIEleuPheGlyGlyGluAspValArgValAlaAsnLeuLysProArgIIEser	680
Db	2062	CAGATGATTCATTTTGGGAGAGAGATGTGCGAGTGGCTAATTTGAAACCAAGAAATCTCC	2121
QY	681	PheAsnPhePheValThrAlaProLysAsnValSerAspIIEIIEProArgThrGluVal	700
Db	2122	TTTAAATTCCTTGTGCACGCACTAAAAATGTGCTGATATCATCAATTCAGAACGTGAAT	2181
QY	701	GluLysAlaIIEArgMetSerArgSerArgIIEAsnAspAlaPheArgLeuAsnAspAsn	720
Db	2182	GAAGAAGGCATCAGGATGTCCCGAGCGCGTATCATGATGTCTTCCGTGATGATCAAC	2241
QY	721	SerLeuGluPheLeuGlyIIEGlnProThrLeuGlyProProAsnGlnProProValSer	740
Db	2242	AGCCTAGAGTTTCTGGGATACAGCAACACATGTGACCTCTAACACACCCCTGTGTTCC	2301
QY	741	IIErPleuIIEValPheGlyValAlaMetGlyValIIEValIValGlyIIEValIIEleu	760
Db	2302	ATATGCTGATGTGTTTGTGGAGTTGTGATGGGAGTGAATGAGTGGCATTTGCTACTCTG	2361
QY	761	IIEPheThrGlyIIEArgAspArgLysLysLysAsnLysAlaArgSerGCGlyGluAsnPro	780
Db	2362	ATCTTCACCTGGGATCAGAGATCGAAGAAAGAAAAATTAAGCAAGAGTGAGAAAAATCT	2421
QY	781	TyrAlaSerIIEAspIIEserIIElysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp	800
Db	2422	TATGCTTCATTCGATTAATTAAGCAAAAGAAAAATTAATCCAGGATTCCAAACACATGATGAT	2481
QY	801	ValGlnThrSerPhe 805	
Db	2482	GTTTCAGACCTCTT 2496	
RESULT	10		
AX431513			
LOCUS	AX431513	3396 bp	DNA linear PAT 28-JUN-2002
DEFINITION	Sequence 1 from Patent WO0239997.		
ACCESSION	AX431513		
VERSION	AX431513.1	GI:21656345	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1		

QY	161	ArgLeuTrpAlaIatPbLuseTrrPArSergLvaLclYlsgLIneuarPrOLEuYr	180
Db	562	AGGCTCTGGGCTTGAGAAAGCTGGAGATCTGAGGTCGGCAAGAGACGTGAGCCATTATAT	621
OY	181	GIuGIuYrValValLeuLysAsnGCIuMeTAlaArgAlaasnHsIrrYrGuaSPYrGly	200
Db	622	GAAGAGTATGTGTCCTTGAAAAATGAGATGGCAAGACCAATCATATATGAGCATGTGG	681
OY	201	AspYrTrpArgGlyAspYrYrGluValAsnGlyValAspGlyTyrrAspYrSerArgly	220
Db	662	GATTATTGGAGAGAGACTATGAAGTAAAGGGGTAGATGTCATGACTACACACCCGCGC	741
OY	221	GIuLeuIleGIuAspValGIuHsIsthPhaGIuGIuLysProLeuYrGIuHsIleu	240
Db	742	CAGTGGATGSAAGATGTGGAAACATACCTTTGAAAGAGTTAAACCATTAATATGAACTCT	801
OY	241	HisAlaYrValArGaLalYsLeuMetAsnAlaYrProSerYrIleSerProIleGly	260
Db	802	CATCCCTATGTGAGGGCAAAAGTTGATGAAGCCCTATCCCTCATATGCTCAATGTGA	861
OY	261	CysLeuProAlaHsIleLeuGlyAspMetTrpGlyArgPheTrpHrsuLeuYrSer	280
Db	862	TGCCCTCCCTGCTCTTTCCTTTCCTTGATATGTGGGGTAGATTTTGGCAAAATGTGACTCT	921
OY	281	LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln	300
Db	922	TTTGACATTCCTCTTTGGACGAAACCAACATGATGTTACTGATCATGGTGGACGCG	981
OY	301	AlaIrrAspAlaGlnArgIlePheLysGluAlaGlyLysPhePheValSerValGlyLeu	320
Db	982	GCCCTGGATGCACAGAAATATTCAGAGAGCCGGAAGACTCTTGTATCTTGTCGCTCT	1041
OY	321	ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln	340
Db	1042	CTTAAATGTGACTCAAGAGATCTTGGGAAATTTCCATGCTTAACGGACCCAGAAATTTTCAG	1101
OY	341	LysAlaValLysHsProThrAlaIrrAspAspLeuGlyLysGlyAspPheArgIleLeuMet	360
Db	1102	AAAGCACTCTGCCATCCCAACAGCTTTGGAGCTGGGGAAAGGGGACCTCGAGATCCTTATG	1161
OY	361	CysThrLysValThrMetAspAspPheLeuThrAlaHsIleGluMetLysIleGln	380
Db	1162	TGCACAAAGGTGACAAATGAGACGACTTCGAGACGCTCATCAGATGGGCAATTCAG	1221
OY	381	TyrAspMetLarYrAlaIalagInProPheLeuLeuArgAsnGlyValAsnGlyPhe	400
Db	1222	TATGATATGGCAATCTGTGCACAACTTTTCTGTAAAGAAAGAGCATATGAAGATTC	1281
OY	401	HisGluValValGlyGluIleMetSerLeuSerAlaIalThrProLysHsIleuYrSer	420
Db	1282	CATGAACGTGTGGGAAATCATGTACTTCTTGCGACGCACACCTTACGATTTAAATCC	1341
OY	421	IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu	440
Db	1342	ATTGGCTCTTCGACCCGCAATTTTCAAGAAGCAATGAAGACGAATAACTCTCTGCTC	1401
OY	441	LysGluAlaLeuThrIleValGlyHrsLeuProPheThrTyrrMetLeuGluLysTrpArg	460
Db	1402	AAACAGACACTCGAATGTGGAGCTGCGCAATTTACTTACATGTTGAAGAGGAGCG	1461
OY	461	TrpMetValPheLysGlyLuuLeuProLysAspGlnTrpMetLysLysTrpGluMet	480
Db	1462	TGATATGCTCTTAAAGGGAATTCOCMAAGACGAGTGAATGAANAATGTGTGGAGATG	1521
OY	481	LysArgGluLuuLuuLuuLuuValGluProValProHsAspGluThrTyrrCysAspPro	500
Db	1522	AAAGCGAATATAGTGGGTGTGGAACTGTGCCCATGTATGAAGACTTACTTGACCCC	1581
OY	501	AlaSerLeuPheHisValSerAsnAspYrSerPheIleArgYrTyrrThrArgThrLeu	520
Db	1582	GCATCTCTGTCCAGTGTCTTAATATATTCATTCATTCATATTCACACAGAACCTTT	1641

Db 344 CAATGCTTCACATACAGAAATTCAGAAATTCACAGCTCAAGCTTCAGTCGACGGCTT 403
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerIysArgLeuAsnThrIleLeu 120
 Db 404 CAGCAAAATGGGTCTTCAGTCTCTCTCAGAAAGACAGAGCAAAAGCGTTGAACACAAATTCCTA 463
 QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
 Db 464 AATCAAAAGACACCATCTACAGTACAGTAAAGATTGTGTACCCAGATATCCACAGAA 523
 QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
 Db 524 TGCTATTACTGAAACAGGTTTGAAATGAAATATGCAAAACGTTTAACTACATCAATGAG 583
 QY 161 ArgLeuThrPalatrrpGluSerTrrpArgSerGluValGlyLysGlnLeuArgProLeuTyr 180
 Db 584 AGGCTCGGGCTGGGAAAGCTGAGAGTCTGAGGTGGCAGACAGCTGAGGCGCAATTAAT 643
 QY 181 GluGluTyrValValleuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200
 Db 644 GAAAGATATGCTTGGTGAAGAAATGAGATGGCAAGAGCAAAATCATTAATGAGCATATGG 703
 QY 201 AspTyrTrrpArgLysAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgly 220
 Db 704 GATTATTTGAGAGAGACATATGAAATAGGGGTATGATGGCTATGACTACAGCCGCGC 763
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240
 Db 764 CAGTTATTTGAAATGTGAAATACCTTTGAAGAGATTAACCATTTATATGAAATCTT 823
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
 Db 824 CAGCTTATGTAGGGCAAGATTGATGAAATGCTTATTCCTTATATGACCAATTTGA 883
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrrpGlyArgPheTrrpThrAsnLeuTyrSer 280
 Db 884 TGCTCTCCCTCATTTGCTGTGATATGTTGGGTATATTTTGACAAATCTGTACTCT 943
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
 Db 944 TTGACAGCTTCCCTTTGGACAGAAACCAATAGATTTACTGATGCAATGGTGGACAG 1003
 QY 301 AlaTrrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320
 Db 1004 GCCTGGGATGACACAGATATTCAGAGAGGCCAGAAATCTTGTGATCTGTTGGCTT 1063
 QY 321 ProAsnMetThrGlnGlyPheTrrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
 Db 1064 CCTAATATGACTCAAGGATCTGGGAAATTCATGCTAACGGACCCAGAAATGTTTCA 1123
 QY 341 LysAlaValCysHisProThrAlaTrrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
 Db 1124 AAAGCAGCTCCCATCCACAGCTTTGGGACCTGGGAAAGGCGCATTCAGAGATCTTAA 1183
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGlnLeuMetGlyHisIleGln 380
 Db 1184 TGCACAAAAGGAGCAATGAGAGCTTCTGACAGCTCATCATAGATGGGCGCATATCCAG 1243
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
 Db 1244 TATGATATGGCATATGCTGCAACATCTTCTCTAAGAAATGAGACCTATGACAGATTC 1303
 QY 401 HisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 Db 1304 CATAAACCTGTTGGGAAATCATGTCATCTTTCGACGCCACACCTTAAAGATTTAA 1363
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGlyThrGluIleAsnPheLeuLeu 440
 Db 1364 ATTTGCTTTCTGACCCCGATTTTCAAGAGAAAGAAAGAAATTAACCTTCCCTGCTC 1423
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrrpArg 460
 Db 1424 AAACAAGCACTCAGATTTGTTGGAGCTTGCCATTTACTTACATGTATTAGAAAGTGGAG 1483

QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrrpMetLysTyrTrrpGluMet 480
 Db 1484 TGGATGGCTTTTAAAGGGAAATTTCCCAAGACAGTGGATGAAAGAAAGTGGGAGATG 1543
 QY 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500
 Db 1544 AAGGAGAGATAGTGGGGTGGTGGAGAACCTGTCCCATATGATTAACATATCTGTACCC 1603
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520
 Db 1604 GCATCTCTGTCCATGTTTGAATGATATCTCATTCATTCGATATACAGAGACCCCTT 1663
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHis 540
 Db 1664 TACCAATTCAGTTTCAAGAGACATTTGTCAAGCAGCTTAACATGAAAGGCCCTTGGAC 1723
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560
 Db 1724 AAATGTGACATCTCAAACTCTACAGAACTGGACAGAAACCTGTCAATATGCTGAGCTT 1783
 QY 561 GlyLysSerGluProTrrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsn 580
 Db 1784 GGAATATCAAGAACCTCGAACCTCAGCATTTGAAATGTTGTATGAGCAAGAAAGATGAT 1843
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrrpLeuLysAspGlnAsnLys 600
 Db 1844 GTAAGGCCACTGCTCAACTACTTTCAGCCCTTATTTACTGGCTGAAGACCAACAG 1903
 QY 601 AsnSerPheValGlyTrrpSerThrAspTrrpSerProTyrAlaAspGlnSerIleLysVal 620
 Db 1904 AATTTCTTTTGTGGATGAGTACCGACCTGAGATCCATATGACCAAAAGCATCAAGTG 1963
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrrpAsnAspAsnGluMet 640
 Db 1964 AAGATTAAGCCTTAAATTCAGCTCTTGAGATATTAAGCATATGATGAAACGACAAATGA 2023
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLysValLysAsn 660
 Db 2024 TACCTGTTCGATCAGTTCGTTGCAATGCTATGAGGAGCTACTTTTAAAAAGTAAAAAT 2083
 QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680
 Db 2084 CAGATGATTTCTTTTGGGAGAGAGATGTGCGAGTGGCTAATTTGAAACCAAGAAATCTCC 2143
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700
 Db 2144 TTTAATTTCTTGTACAGCTACCTAAAAATGTGTCTGATTCATTTCTAGAACTGAAGTT 2203
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
 Db 2204 GAAAAGGCATCAGAGATGTCCCGGAGCGGTATCAATGATGCTTCCGTGTAATGACAAC 2263
 QY 721 SerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
 Db 2264 AGCTTAGAGTTTCTGGGATATACGCCAACCTTGAGACCTCTACACAGCCCTGTGTTTC 2323
 QY 741 IleTrrpLeuIleValPheGlyValAlaMetGlyValIleValValGlyIleValIleLeu 760
 Db 2324 ATATGGCTGATGTTTGTGAGTTGTGATGGAGATATGTTGGCTATGTCATCTCG 2383
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780
 Db 2384 ATCTTCACTGGATCAGAGATCGGAAGAAAGAAATTAAGCAAGAAAGTGGAAATATCCT 2443
 QY 781 TyrAlaSerIleAspIleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAsp 800
 Db 2444 TATGCTCTCAATGATATTTGCAAAAGGAAATTAATCCAGATTTCCAAAACACTGATGAT 2503
 QY 801 ValGlnThrSerPhe 805
 Db 2504 GTTCAGACCTCTCTT 2518

RESULT 12
 AB046569
 LOCUS
 DEFINITION Homo sapiens ace2 mRNA, complete cds.
 ACCESSION AB046569
 VERSION AB046569.1 GI:13516971
 KEYWORDS
 SOURCE Homo sapiens adult cDNA to mRNA, clone_lib:Sugano human cDNA
 ORGANISM
 Homo sapiens
 Library Kala clone:kala4505.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 (sites)
 2 (bases 1 to 2599)
 AUTHORS Suzuki, Y., Matanabe, M. and Sugano, S.
 TITLE Cloning, expression analysis and chromosomal localization of a novel ACE like enzyme
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2599)
 AUTHORS Komatsu, T., Suzuki, Y. and Sugano, S.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUL-2000) Takami Komatsu, the Institute of Medical Science, Virology; 4-6-1, Minato-ku, Tokyo 108-8639, Japan
 (E-mail: komatsu@iims.u-tokyo.ac.jp, Tel:81-3-5449-5283(ex.75283), Fax:81-3-5449-5416)
 FEATURES
 source
 1. 2599
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="Xp22"
 /clone="kala4505"
 /clone_lib="Sugano human cDNA library kala"
 /dev_stage="adult"
 /country="Japan"
 /note="Vector: pME18S FL3, lab host: Top10"
 1. 2599
 /gene="ace2"
 55. 2472
 /gene="ace2"
 /codon_start=1
 /product="ACE2"
 /protein_id="BAB40370.1"
 /db_xref="GI:13516972"
 /translation="MSSSWLLSLVAVTAASSTEEQAKFLDKFNHAEADLFYSS
 LASNNVNTTENNQNNMNAAGDKMSAFKESQSLAOMPLLEIGNLVKQLQALQ
 NGSSVLESDSKRLNTILNTMSTIYSTGVKCPNDPOECLLEPGNETMAISLDYNE
 RLWAMESWRESEVKOLRPLEETVYVKNEMANRNEYEDVDYNGDVGVDYS
 RGQILIEVETHEFEIKRPLEYEHIAVYRAKLMANRPEYISPTICLPAHLIGDMGSEMT
 NLVSLVPEFGCKNIDVTAMVDQADQRIKREAKRFVSGLPMTQGEWNSMLI
 DPGNVKAVCHPTAMPDLGKDFRLIMCTVTYMDPDLTAHHEHGIQYDAVAPFL
 RGNANGFEHVAEIMSLAATPKHLKSLISPDQENENETFLKQALTIYGLT
 PFTYMEKRMWVPEKIGIPKDMKKRMEKREIVGVPEVPHDEYCPALFVSN
 DYSFIRYITLLOFOQELCOAAKHGEPHICDISNSTEAGOKLFNMLRGKSEPV
 TLALENVGAKNMNVRLNVEPFLTWLKDQKNSFYVSGSDMSFYADOSTKXVLSL
 KSALEGRAYEMNDENYLFERSVAVAROYFLVKQMLLGEEDYRYANLKRISFN
 FETVAPAKNSDILIPREVERKAIRMSRIDARLNDNSLEFLGIQPTLGPNOQPS
 IMLIVFGVGVIVGIVILIFTGIRDKKKRKRSGENPYASIDISKENNPQNT
 DVQTFSE"
 BASE COUNT 808 a 503 c 590 g 698 t
 ORIGIN
 Alignment Scores:
 pred. No.: 0 Length: 2599
 Score: 4288.00 Matches: 804
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.88% Mismatches: 0
 Query Match: 99.93% Indels: 0
 DB: 9 Gaps: 0
 US-09-978-385-2 (1-805) x AB046569 (1-2599)
 QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 55 ATGTCAGCTCTTCCCTGGCTCTTCTCAGCCTTGTCTGATACGTGCTGATCCACC 114
 QY 21 IleglglglAlaAlaLysThrPheLeuAspLysPheAsnHisglAlaGluAspLeuPhe 40
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 115 ATTGAGGAACAGGCCAACACATTTTGTGACAAAGTTTACACAGCAAGCCGAGACCTGTC 174
 QY 41 TTYGlnSerSerLeuAlaSerTrpAsnThrAsnHisThrAsnHisThrGlnAsnValGln 60
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 175 TATCAAAAGTTCCTGCTTCTTGGATATATACACCAATATTTACTGAAAGAAATGTCACAA 234
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAla 80
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 235 AACATGAATTAACGCTGGGACAAATGCTCTCTTTTAAAGCAAGCTCCACACTGGCC 294
 QY 81 GlnMetTrpProLeuGlnGlnLeuGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 295 CAATGATATCCACTACAGAAATTAAGATTTCTACACAGCAAGCTTCAAGCTGACGCTCTT 354
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrLleLeu 120
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 355 CACCAAAATGGGCTTCAAGCTCTCAGAACAGCAAGCAAGCAAGCTTGAACCAATTTCTA 414
 QY 121 AsnThrMetSerThrLleTrpSerThrGlyLysValLysAsnProAspAsnProGlnGlu 140
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 415 AATACATGAGCACCATCTACAGTCTGAGAAAGTTTGTAACCCAGATTAATCCACAAAGA 474
 QY 141 CysLeuLeuLeuGluProGlyLeuAsnGlnLleMetAlaAsnSerLeuAspTrpAsnGlu 160
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 475 TGCTTATTTACTTGAACACAGTTTGAATGAATATGCAACAGTTTACGTACACATGAG 534
 QY 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTrp 180
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 535 AGGCTCGGCTTGGGAAAGCTGAGATCTGAGTGGAGTGGCAACACCTGAGCCATTAAT 594
 QY 181 GlnGluTrpValValLeuLysAsnGlnMetAlaArgAlaAsnHisTrpGluAspTrpGly 200
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 595 GAAAGATATGCTGCTTGAATAATGAGATGGCAAGCAACATCATTAAGAGCATATGAG 654
 QY 201 AspTrpTrpArgGlyAspTrpGluValAsnGlyValAspGlyLysAspTrpSerArgGly 220
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 655 GATATTTGAGAGAGAGCATATGAAGTAATAGGGGTGATGAGCTATGACTACAGCCGCCG 714
 QY 221 GlnLeuLeuGluAspValGlnHisThrPheGlnGluLleLysProLeuTrpGlnHisLeu 240
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 715 CAGTTGATTAAGATGTGGAAACATACCTTTGAAGAGATTAAACCATTAATGAACATCTT 774
 QY 241 HisAlaTrpValAlaArgAlaLysLeuMetAsnAlaTrpProSerTrpLleSerProIlegly 260
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 775 CATGCTATGTGAGAGGCAAAAGTTGATGAATGCTATCTCTTAATATCAAGTCCAAATTGGA 834
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpHisAsnLeuTrpSer 280
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 835 TGCCCTCCCTGCTCATTTGCTTGTGATATATGGGGATGATTTTGGCAAAATCGTACTCT 894
 QY 281 LeuThrValProPheGlyGlnLysProAsnLleAspValThrAspAlaMetValAspGln 300
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 895 TTGACAGTTCCCTTTGGACAGAAACCAACATAGATTTACTATGAAAGGGGAGCCAG 954
 QY 301 AlaTrpAspAlaGlnArgLlePheLysGlnAlaLysPhePheValSerValGlyLeu 320
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 955 GCCTGGAGATGACACAGAAATATTCAGAGAGGCGAGAAAGTTCTTTGTATCTGTGGCTCTT 1014
 QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1015 CCTAATATGACTCAAGGATTTCTGGAAATATTCATGCTAACGAGCCAGAAATGTTTCAG 1074
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgLleLeuMet 360
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1075 AAAGCAGCTGCCATCCACACAGCTTGGGACCTGGGAAAGCCGCTTCAGATCTTATG 1134
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyHisLleGln 380
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1135 TGCACAAAGGTACAAATGAGACGATTTCTGACAGCTCATATGAGATGGGCAATATTCAG 1194

QY 381 TyrAspMetAlaTyrAlaIaInProPheLeuLeuArgSngLYAlaasnGluGlyPhe 400
 |||||
 Db 1195 TATGATATGGCATATCTCTGCACACCTTTTCTTCTGCTAGAGAAAGGAGCTATATAGAGATTC 1254
 QY 401 HisGluAlaValGlyIuileMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 |||||
 Db 1255 CATGAGAGCTGTGGGGAATCATGTCTTCTGCAGGCCACACCTAAGATTTAAATCC 1314
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspSngLThrGluIleAsnPheLeuLeu 440
 |||||
 Db 1315 ATTTGGCTTCTGTCCACCATTTTCAAGAAAGCAATGAAACAGAAATTAACCTCTGCTC 1374
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460
 |||||
 Db 1375 AACAAGACCTCAGCATTTCTGGACCTGCGCATTTACTTACATTTTGAGAGTGGAGG 1434
 QY 461 TrpMetValPheLysGlyIuileProLysAspGlnTrpMetLysLysTrpPheLysMet 480
 |||||
 Db 1435 TGGATGCTCTTAAAGGGGAATTCCTCAAGACAGTGGATGAAAAAGTGGGAGATG 1494
 QY 481 LysArgGluIleValGlyValGluLupProValProHisAspGluThrTyrCysAspPro 500
 |||||
 Db 1495 AAGCGAGAGATAGTTGGGTGGTGGACCTGCCCCATGATGAAACATACCTGTCACCC 1554
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520
 |||||
 Db 1555 GCATCTCTGTCATGTTCTTAATATTAATCACTTCACTTCATTAACACAGGACCTT 1614
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGlySngAlaAlaLysHisGluGlyProLeuHis 540
 |||||
 Db 1615 TACCAATTCAGATTTCAGAGACCTTTGTCAAGCAGCTAAACAGAGAGGCCCTGTGAC 1674
 QY 541 LysCysAspLysSerAsnSerThrGluAlaGlyGluLysLeuPheAsnMetLeuArgLeu 560
 |||||
 Db 1675 AATATGACATCTCAAACTCTACAGAACCTGACAGAACGTTCAATATCTGAGGCTT 1734
 QY 561 GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValAlaGlyAlaLysAsnMetAsn 580
 |||||
 Db 1735 GGAATATCAGAAACCTGGACCTTACATTTGGAATATGTTGAGAGCAAAAGCATGAT 1794
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpPheLysAspGlnAsnLys 600
 |||||
 Db 1795 GTAAGGCCACTGCTCACTACTTTGAGCCCTTATTTACCTGGCTGGAAGACAGAACAG 1854
 QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620
 |||||
 Db 1855 AATCTTTTGTGGATGAGTACCGACTGAGTGCATATGCAAGCAAGCATCAAGATG 1914
 QY 621 ArgLysSerLeuLysSerAlaLeuGlyLysPylsAlaTyrGluTyrPasnAspAsnGluMet 640
 |||||
 Db 1915 AGGATTAACCTAAATATCAGCTCTTGAGATAGACATATGAATGAAGACAGACATGAATG 1974
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660
 |||||
 Db 1975 TACCTGTTCCATCATCTGTTGCAATATGCTATGAGGAGCTTTTAAAGGTAAAAAAT 2034
 QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgLysSer 680
 |||||
 Db 2035 CAGATGATTTCTTTTGGGAGAGAGATGTGGAGTGGCTAATTTGAAACAGAAATCTCC 2094
 QY 681 PheAsnProPheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700
 |||||
 Db 2095 TTTAATTTCTTTCACACGCACTTAAATATGCTCATATATCATTTCTAGAACTGAAGTT 2154
 QY 701 GluLysAlaIleArgMetSerArgSerArgLysAsnAspAlaPheArgLeuAsnAspAsn 720
 |||||
 Db 2155 GAAAAGCCATCAGAGATGTCCCGAGCCGATATCATATGCTTCCCTCTGAATGAACAAC 2214
 QY 721 SerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
 |||||
 Db 2215 AGCTAGAGTTTCTGGGATACAGCAAACTTGGACCTCTAATCAAGACCCCTGTTTCC 2274

QY 741 IleTrpLeuIleValPheGlyValAlaMetClyValIleValValGlyIleValIleLeu 760
 |||||
 Db 2275 ATATGGCTGATTTGTTTGGAGTTGTGATGGAGTGTAGTGTGATTCATTCATCTCG 2334
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysPro 780
 |||||
 Db 2335 ATCTTCACTGGGATCAGAGATCGGAGAGAAAAATTAACAGCAAGATGGAGAAATCT 2394
 QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAsp 800
 |||||
 Db 2395 TATGCTTCATCATGATATTAGCAAGAGAAAAATATCATCAGATTTCAAAACATGATGAT 2454
 QY 801 ValGlnThrSerPhe 805
 |||||
 Db 2455 GTTCAGACCTCTT 2469
 RESULT 13
 HSM800880
 LOCUS
 DEFINITION Homo sapiens mRNA; cdna DKFZp434A014 (from clone DKFZp434A014);
 partial cds.
 ACCESSION AL110224
 VERSION AL110224.1 GI:5817159
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens.
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 (bases 1 to 3341)
 Wambutt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
 JOURNAL Submitted (15-AUG-1999) MIPs, Am Klopfersplitz 18a, D-82152
 Martinsried, GERMANY
 COMMENT
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by AGOWA (Berlin/Germany) within the cdna sequencing
 consortium of the German Genome Project.
 This clone (DKFZp434A014) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at http://www.mips.biochem.mpg.de/proj/cdna/.
 FEATURES
 source
 1. 3341
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp434A014"
 /issue_type="testis"
 /clone_lib="434 (synonym: htes3). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 1. 2417
 /gene="DKFZp434A014"
 <1. 2417
 /gene="DKFZp434A014"
 /note="strong similarity to peptidyl-dipeptidase A"
 /protein_id="CAB53682.1"
 /db_xref="GI:5817160"
 /translation="SSSSWLLSLVAVTAHSTTEDEAKFTLDFKFNHEADLFFQSSL
 ASMNNTNTEENVONMNNAGKNSAFKEBSTLAQMTLOEIOINLTKILOALQON
 GSSVLEDEKSKRLNTLNMSTIYSGTGVCPNDNQBELLPGINEMANSLDYNER
 LMAWMSSEVCKOIRPLYEEVYVKNEMARNRHEDYGDYWRGDIYVNGVDYISR
 GQLIDVETHEPEIKPLVEHHAIVRAKLMAVPSYISPCILPAHLIGDWGRFWN
 LYSILVPGQKRNIDVTAMVDQADADQRIKREKEFEVSGVGLNMTQGEWENSLMD
 PGNVKAACHPTAMDIGKDFRIIMCTGVTDPLTAHHEMGILOYMYAAGPFLR
 NGANGFEHAYGEIWSLAATPKHLKSGLSPPDEENETETINELKQALATVGTLP
 FYMLEKRWVVFKEIPIKDDMMKKMKMKREIVGYVPPVPHDETCDDPASLPHVSD
 YSFTIYVRLTXQPOFOALQAKHBEPLKCDISNTEAGQLEFNLRLGKSEPT
 LALENVCAKMNANRPLNIFEPLETLKLDONKNSFGWSTWDSPYADQSLKVISLK
 SALGKAYEMNDENMYLRSSVAYAMROYFLKVNOMTLFGEEDVVRANLKRISFNR
 FVLAENKVSIDIIPTEVEKAIKRSRSLINDAFRLNDNSLEFLGIQPTLGPNQPPVSI

WLIVGVGMIVIVGIVILLFTGIRDRKKKKKARSGENPYASIDISGENNPGFQND
 DVQSF*
 polyA_signal 3266..3271
 polyA_site 3281
 BASE COUNT 1058 a 630 c 740 g 913 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3341
 Score: 4281.00 Matches: 803
 Percent Similarity: 99.88% Conservative: 0
 Best Local Similarity: 99.88% Mismatches: 1
 Query Match: 99.77% Indels: 0
 DB: 9 Gaps: 0

US-09-978-385-2 (1-805) x HSM800880 (1-3341)

QY 2 SerSerSerSerTrpleuLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThrIle 21
 |||||
 DB 3 TCAGCTCTTCTGCTGCTCTTCTCAGCCCTGTTGGCTGTAAGCTGCTCAGCCACAT 62
 QY 22 GlnGlnGlnAlaIleThrPheLeuSplysPheAsnHisGlnAlaGlnSpleuPheTyr 41
 |||||
 DB 63 GAGGACAGCGCAAGACATTTTGGCAAGTTTAAACCAAGCCGCAAGCCTGTTCTAT 122
 QY 42 GlnSerSerLeuAlaSerTrpAsnTrpAsnThrAsnIleThrGlnGlnAsnValGlnSn 61
 |||||
 DB 123 CAAAGTCACTGCTTCTGGAATTTACACCAATATTACTGAAGACATGTCACAAAC 182
 QY 62 MetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAlaGln 81
 |||||
 DB 183 ATGATTAATAGCTGGGAGCAATGCTGCTTTTAAAGAACAGCTCACACTTGCCCA 242
 QY 82 MetTrpProLeuGlnIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeuGln 101
 |||||
 DB 243 ATGTATCCACTACAAAGAAATTCAGATTCACAGTCACAGCTTCAAGCTGAGCTTCAG 302
 QY 102 GlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThrIleLeuAsn 121
 |||||
 DB 303 CAAATAGGCTCTTCACTGCTCTCAGAAAGCAAGACAAACGGTTGACACATTTCTAAT 362
 QY 122 ThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlyCys 141
 |||||
 DB 363 ACAATGAGCACCACTTACAGTACTGGAAGATTGTGAACCCAGATTAATCCAAAGATGTC 422
 QY 142 LeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspLysArgLeuGln 161
 |||||
 DB 423 TTATTACTTGAAACCAAGTTTGAATGAATTAATGCAAAACAGTTTACATCAATGAGAG 482
 QY 162 LeuTrpAlaTrpGlnSerTrpArgSerGlnValGlyLysGlnLeuArgProLeuTyrGln 181
 |||||
 DB 483 CTCTGGGCTTGGGAAAGCTGAGATCTGAGTGGCAACAGCTGAGGCGCATTAATATGA 542
 QY 182 GlnTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnSplysArgLeu 201
 |||||
 DB 543 GAGTATGCTGCTTGAAGAAATGAGATGAGCAAGCAAAATCATTAATGAGACATATGGGAT 602
 QY 202 TyrTrpArgGlyAspTyrGlnValAsnGlnValAspGlyTyrAspLysSerArgLeuGln 221
 |||||
 DB 603 TATTTGAGAGAGACTATGAATTAATGGGATGATGGCTATGACTACAGCCGCGCCAG 662
 QY 222 LeuIleGlnAspValGlnHisThrPheGlnGlnIleLysProLeuTyrGlnHisLeuHis 241
 |||||
 DB 663 TTGATTCAGACATGTCGACACTTCTGGAAGATTAACCATTAATATGACATCTTCAT 722
 QY 242 AlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGlyCys 261
 |||||
 DB 723 GCTATGCTGAGGCAAAAGTTGATGAATGCTATCTCTATATCAGTCCCATTTGATGAC 782
 QY 262 LeuProAlaHisLeuLeuGlnGlyAspMetTrpGlyArgPheTrpHisLeuTyrSerLeu 281
 |||||
 DB 783 CTCCTGCTCATTTGCTGCTGATATGTGGGTAGATTTTGGCAAAATCTGTACTCTTGG 842

QY 282 ThrValProPheGlnGlnLysProAsnIleAspValThrAspAlaMetValAspGlnAla 301
 |||||
 DB 843 ACAGTTCCTTTGGACGAGAAACCAACATATGATTATCATATGATGATGGTGGACGAGCC 902
 QY 302 TrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerValGlyLeuPro 321
 |||||
 DB 903 TGGGATCAGACAGAAATATTCAGAGAGCCGCAAGATCTTTGATCTGTTGGCTTCTCT 962
 QY 322 AsnMetThrGlnGlnPheTrpGlnAsnSerMetLeuThrAspProGlnAsnValGlnLys 341
 |||||
 DB 963 AATATGACTCAAGGATTTCTGGAAATTTCCATGCTAAGGACCCAGAAATGTTCCAGAAA 1022
 QY 342 AlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMetCys 361
 |||||
 DB 1023 GCAGTGGCCATCCACAGCTTGGAGCTGGGGAAGGCGACCTTCAAGATCTTATATGTGC 1082
 QY 362 ThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlnHisIleGlnTyr 381
 |||||
 DB 1083 ACAAGGTGACAAATGAGACGACTTCTGACAGCTCATCATGATGAGTGGGCAATTCAGTAT 1142
 QY 382 AspMetAlaTyrAlaIleAlaGlnProPheLeuLeuArgAsnGlnValAsnGlnGlyPheHis 401
 |||||
 DB 1143 GATATGGCATATGCTGACACAACTTTTCTGCTAAGAAATGAGACTAATGAAGATTCAT 1202
 QY 402 GlnAlaValGlyGlnIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSerIle 421
 |||||
 DB 1203 GAAGCTGTTGGGAAATCATGTCACCTTCTGACCCACCTTAAGCATTAATAATCATAT 1262
 QY 422 GlyLeuLeuSerProAspPheGlnGlnLysAspAsnGlnIleThrGlnLysAsnPheLeuLys 441
 |||||
 DB 1263 GGTCTTGTGTCACCCGATTTTCAAGAAACAAATGAACAAATGAACAAATTTCTGCTCAA 1322
 QY 442 GlnAlaLeuThrIleValGlyThrLeuProPheThrIleMetLeuGlnLysTrpArgTyr 461
 |||||
 DB 1323 CAAGCAGTCACAGATTTGGGAGCTGCTGCCATTTACTTACATGTTAGGAAGTGGAGCTGG 1382
 QY 462 MetValPheLysGlyGlnIleProLysAspGlnTrpMetLysLysTrpArgLeuMetLys 481
 |||||
 DB 1383 ATGCTCTTAAAGGGAATCCCAAGACAGAGTGATGAAGAAAGCTGTTGGGAGATGAG 1442
 QY 482 ArgGlnIleValGlyValAlaGlnProValProHisAspGlnIleThrTyrCysAspProAla 501
 |||||
 DB 1443 CGAGAGATATGGGAGTGGTGAACCTGCTGCCATGATGAACATCTGAGACCCGCA 1502
 QY 502 SerLeuPheHisValSerAsnAspLysSerPheIleArgTyrTrpThrArgThrLeuTyr 521
 |||||
 DB 1503 TCTGCTTCATGTTTCAATGATTAATCAATTCATTCATGATTAATCAACAGACCTTTAC 1562
 QY 522 GlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHisLys 541
 |||||
 DB 1563 CAATTCAGATTCAAGAGCACTTGTGCAAGCACCTTAACATGAAAGCCCTTGCACAAA 1622
 QY 542 CysAspIleSerAsnSerThrGlnAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGly 561
 |||||
 DB 1623 TGTGACATCTCAAACTCTACAGAAAGCTGGAGAGAACTGTTCAATATGCTGAGGCTTGA 1682
 QY 562 LysSerGlnProTrpThrLeuAlaLeuGlnAsnValAlaGlyAlaLysAsnMetAsnVal 581
 |||||
 DB 1683 AAATAGAAACCTGGACCTGACCTGATGAAATATGTTAGAGAGCAAGAACATGAATGTA 1742
 QY 582 ArgProLeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuLysAspGlnAsnLysAsn 601
 |||||
 DB 1743 AGGCCACGCTCAACTACTTGTGAGCCCTTATTTACTGCTGCGCAAGACCGAAGAACAAAT 1802
 QY 602 SerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysValArg 621
 |||||
 DB 1803 TCTTTTGGGATGGAGTACGCACTGAGTCAATATGCAAGCCAAAGCATCAAAAGTACG 1862
 QY 622 IleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGlnTyrPheAsnAspAsnGlnMetTyr 641
 |||||
 DB 1863 ATTAAGCTTAATAATACGCTTGGAGATTAACCATATGAAATGAGACGCAATGAATGTAC 1922
 QY 642 LeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGln 661

```

|||||
Db 1923 CCGTCCGATCATCTGTTCCATATGCTATGAGGCGAGTACTTTTAAAGTAAATAATCAG 1982
QY 662 MetileLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgLleSerPhe 681
Db 1983 ATGATCTCTTTTGGGAGAGAGATGGAGTGGCTAAATTTGAAACCAAGAAATCTCTTT 2042
QY 682 AsnPheValThrAlaProLysAsnValSerAspLleLeuProArgThrGluValGlu 701
Db 2043 AATTTCTTGTGCTGACCTCAATAAATGCTGATATCATCTCCAGAACTAAATTGAA 2102
QY 702 LysAlaIleArgMetSerArgSerArgLleAsnAspAlaPheArgLeuAsnAspAsnSer 721
Db 2103 AAGGCATCAGAGATGTCGGAGCCGCTATCAATGATGCTTCCGTCAATGACAACAGC 2162
QY 722 LeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerLle 741
Db 2163 CTAGAGTTTCTGGGATACAGCCAACTGGACCTCTTACCAACCCCTGTTTCCATA 2222
QY 742 TrpLeuIleValPheGlyValValMetGlyValIleValIleValIleValIle 761
Db 2223 TGGCTGATTTGTTTGGAGTTGTGATGGAGATGATGATGATGATGATGATGATGATG 2282
QY 762 PheThrGlyIleArgAspArgLysLysLysLysLysLysLysLysLysLysLysLys 781
Db 2283 TTCACCTGGATCAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2342
QY 782 AlaSerIleAspLleSerLysGlyLysAsnProGlyPheGlnAsnThrAspAspVal 801
Db 2343 GCCTCATCGATATTAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2402
QY 802 GlnThrSerPhe 805
Db 2403 CAGACCTCTCTT 2414

RESULT 14
AX463938 3732 bp DNA linear PAT 16-JUN-2002
LOCUS AX463938
DEFINITION Sequence 71 from Patent WO0140466.
ACCESSION AX463938
VERSION AX463938.1 GI:21898986
KEYWORDS
SOURCE human.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E.,
Gao,W.Q., Gerlitsen,M.E., Goddard,A., Goddard,P.J., Gurney,A.L.,
Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,
Wood,W.L. and Zhang,Z.
Secreted and transmembrane polypeptides and nucleic acids encoding
TITLE
JOURNAL
same
Patent: WO 0140466-A 71 07-JUN-2001;
Genentech Inc. (US)
FEATURES
source 1..3732
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1137 a 722 c 821 g 1052 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3732
Score: 4142.00 Matches: 802
Percent Similarity: 85.33% Conservative: 1
Best Local Similarity: 85.23% Mismatches: 2
Query Match: 96.53% Indels: 138
DB: 6 Gaps: 1
US-09-978-385-2 (1-805) x AX463938 (1-3732)
QY 1 MetSerSerSerThrLeuLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20

```

```

|||||
Db 40 ATGTCACACCTTCTTCTGCTCTCTCAAGCCTTCTTCTGTAAGTCTGCTCACTCCACC 99
QY 21 IleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
Db 100 ATTGAGAGACAGGCGCAAGACATTTTGGACAGTTTAAACCAAGCCAGAGACCTGTC 139
QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsnValGln 60
Db 160 TATCAAAAGTTCACTTGTCTTCTTGAATTAACCCAAATATTACTGAAGAGAAATGTC 219
QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
Db 220 AACATGATTAATGCTGGGAGCAAAATGCTGCTTTTAAAGAGACGTCCACACTTGGC 279
QY 81 GlnMetLysProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
Db 280 CAAATGATTCACACACAAAGAAATTCAGATCTCAGAGTCAAGCTTCAGCTGACGCTCT 339
QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
Db 340 CAGCAAAATGGCTTCAAGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 399
QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
Db 400 AATTCATATGAGCACACCATCTACAGTACGAGAAAGTTTGTAAACCAATATTCACAAAG 459
QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
Db 460 TGCTTATTACTTGAACAGGCTTGAATGAATATGCAACAGTTTACACTCAATGAG 519
QY 161 ArgLeuThrAlaTrpGlnSerTrpArgSerGluValGlyGlnLeuArgProLeuTyr 180
Db 520 AGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGAGTGGCAAGCAGCTGAGCCATTATAT 579
QY 181 GluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200
Db 580 GAAGAGATATGCTCTTGAAGAAATGAGATGGCAAGAGCAAAATCATATATAGAGCTATGG 639
QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
Db 640 GATTATTGGAGAGAGACATATGAATGAATGGGCTGATGCTATGATCAGTCAAGCCGCGC 699
QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlnHisLeu 240
Db 700 CAGTTGATTAAGAAATGTGGAACATACCTTTGAAGAGATTAACCATTAATATAGACATCT 759
QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db 760 CAGGCTATGTGAGGCAAAAGTTGATGAATGCTATCTCTTAATATCAGTCCAAATGGA 819
QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheThrAsnLeuTyrSer 280
Db 820 TGGCTCCCTGCTCATTTGCTTGGTATATGTGGAGATTTTGGACCAATCTGATCTCT 879
QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 880 TTGACAGTTCCTCTTGGACAGAAACCAACATAGATGTTACATGATGCAATGGTGGACAG 939
QY 301 AlaTrpAspAlaGlnArgLlePheLysGluAlaGluLysPhePheValSerValIleLeu 320
Db 940 GCGTGGAGTACAGAGAAATATCAAGAGGCGGAGAAATCTTGTATCTGTGCTCTT 999
QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyValAsnValGln 340
Db 1000 CCTAATATATACACCAAGATATTCGGGAAATTTCACTATACGAGACCCAGAGAAATTTAG 1059
QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1060 AAAGCAGTCTGGCAATCCACAGCTTGGGAGCTGGGAGAGGGGAGACTTCAGATCTTATG 1119
QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380

```

Db	1120	TGCAAGAAGGACATGAGACGACTTCTGACAGCTCATCATGATGAGGAGCATTCAG	1179	Qy	606	TrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleuysValArgIleSerIleuLy	625
Qy	381	TyrAspMetAlaTyrAlaAlaGlnProPheIleuArgAsnGlyAlaAsnGlyPhe	400	Db	2258	TGAGATACCGACTGGAGTCCATATGACAGCCCAAGACATCAAGAGATTAAGCTTAA	2317
Db	1180	TATGATATGCGATATGCTGCAACAACCTTTCTGCTAGAAATGAGCTAATGAAAGATTTC	1239	Qy	625	sSerAlaIleuGlyAspLysAlaTyr-GluTrpAsnAspAsnGluMetTyrLeuPheArgSe	645
Qy	401	HisGluAlaValGlyGluIleMetSerLeuSerIleAlaIleProIleuSHSleuLysSer	420	Db	2318	ATCAGCTCTTGAGATTAAGCATATGAAATGAAACAGCAATGAAATGCTGCTCCGATC	2377
Db	1240	CATTAAGCTGTGGGAAATCATCTCTCTGACCCACCTTAAGCATTTAAATTC	1299	Qy	645	rSerValAlaTyrAlaMetArgGlnTrpPheIleuysValLysAsnGluMetIleLeuPh	665
Qy	421	IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluTrpGluIleAsnIleuLeu	440	Db	2378	ATCTGTTCATATGCTATGAGCAGACTCTTTTAAAGTAAATAATCAGATATTCCTTTT	2437
Db	1300	ATTGGTCTCTGTCACCCGATTTTCAGAGACATGAACAAGAAATAAATCTGCTCCTC	1359	Qy	665	eglyGluGluAspValArgValAlaAsnLeuysProArgIleSerPheAsnIleuPheVa	685
Qy	441	LysGlnAlaIleuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg	460	Db	2438	TGGGAGAGAGATGTCGAGCTGCTATTTGAACCAAGAAATCTCTTAATTTCTTTGT	2497
Db	1360	AAACAGACACTACGATGTGGGACTCTGCAATTTACTTACATGATGAGAAAGTGGAGG	1419	Qy	685	IThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAlaIleAr	705
Qy	461	TrpMetValIlePheLysGlyGluIleProLysAspGlnTrpMetLysTrpTrpGluMet	480	Db	2498	CACCTGACCTAAATAATGTCTGATATCATTTCTAGAACTGAAGTGAATAAGCCATCAG	2557
Db	1420	TGGATGGCTTTAAAGGGGAAATTCACCAAGACAGTGGATGAAAGTGGTGGAGATG	1479	Qy	705	gMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLe	725
Qy	481	LysArgGluIleValGlyValIleGluProValProHisAspGluThrTyrCysAspPro	500	Db	2558	GATGTCCCGAGCGGATGATGATGATGCTTCCGTGAAATGCAACAGCTGAGATTCT	2617
Db	1480	AAGCAGAGATAGTTGGGGTGGTGGAACTGTGCCCATGATGAACATACTGTGACCCC	1539	Qy	725	uglyIleGlnProThrLeuGlyProProAsnGlnProProValSerIleTrpLeuIleVa	745
Qy	501	AlaSerIleuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu	520	Db	2618	GGGAGATACAGCCAAACCTTGACCTCTTAACAGCCCGCTGTTCCATATGCTGATGCT	2677
Db	1540	GCATCTGTCTTCATGTTCTGATGATGATCTCATTCATTCATATTCACAAAGACCTT	1599	Qy	745	IleGlyValAlaMetGlyValIleValAlaGlyIleValIleLeuIlePheThrGlyIle	765
Qy	521	TyrGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaIleAlaIleSHSleuGlyProLeuHis	540	Db	2678	TTTTGGAGTGTGATGGAGTGTGATGCTGTGGCATTCATTCCTCATCTCATCTGCGGAT	2737
Db	1600	TACCATTCCTGATTCAGAAAGACCTTTGTCAAGCAGCTAAACATGAAAGGCCCTGTCAC	1659	Qy	765	eArgAspArgLysLysLysAsnLysAlaIleArgSerGlyGluAsnProTyrAlaSerIleAs	785
Qy	541	LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe-----	555	Db	2738	CAGAGATCGGAGAGAAATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2797
Db	1660	AAATGTGACATCTCAAACTACAGAGACGTGACAGAAAGCTGT- GTAAGAAATACCTCA	1718	Qy	785	PIleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPh	805
Qy	555	-----	555	Db	2798	TATTAGCAAGAGAAATATATCCAGAGATTCCAAAACACTGATGATGATGATGATGAT	2857
Db	1719	AAATGTGAACCTCTCTAGTATTCAGTATTAATTCATTCATTCATTCATTCATTCAT	1778	Qy	805	e 805	
Qy	555	-----	555	Db	2858	T 2858	
Db	1779	ATTCTGTGTTCTAAAGAAATTTTATGGCTCAAAATGTCCTCATTTACAACCAAA	1838	RESULT 15			
Qy	555	-----	555	E39034			
Db	1839	CATTTAATTTGTGTGACAGACAGAACTAGACATACAAATTTGGTGGCCACCTCTT	1898	LOCUS	E39034	2262 bp	DNA linear
Qy	555	-----	555	DEFINITION	MPROT15 polypeptide and MPROT15 polynucleotide.		PAT 18-JUN-2001
Db	1899	TTTCTCCCTATCATTAACACGCTCTCTCTCTGTAATTGGAAGAAAGCGGTTAG	1958	ACCESSION	E39034.1	GI:13017696	
Qy	555	-----	555	VERSION	E39034.1	GI:13017696	
Db	1959	GGTGAATATATCTGTAATATATGATCTTTTCTTATCTGCGAGAGCAAAATTTAGCCAA	2018	KEYWORDS	JP 199318472-A/2.		
Qy	555	-----	555	SOURCE	Homo sapiens.		
Db	2019	GTCAAGAGAGAAACCATATGATCATAGTAAATATATATGATCATCTGGAACCCCTCAA	2078	ORGANISM	Homo sapiens		
Qy	556	-----	556	REFERENCE	1 (bases 1 to 2262)		
Db	2079	AAGGCCCTGACCCCTTTTGTGTAGCATATATGCTGAGGCTTGGAATAATCACAACCC	2138	AUTHORS	Christophe,D.S. and Nicola,B.		
Qy	566	TrpThrLeuAlaLeuGluAsnValIleGlyAlaLysAsnMetAsnValArgProLeuLeu	585	TITLE	MPROT15 polypeptide and MPROT15 polynucleotide		
Db	2139	TGACACCTAGCATTTGGAATAATGTTAGAGAC- AAGAACATGATGTAAGGCGCTGCTC	2197	JOURNAL	Patent: JP 199318472-A 2 24-NOV-1999;		
Qy	586	AsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysSerPheValGly	605	COMMENT	SMITHKLINE BEECHAM CORP PUBLIC LTD CO		
Db	2198	AACTACTTTGAGCCCTTATTTACTGTGGTGAAGACCAAGACAGAAATTTCTTTTGGGA	2257	OS	Homo sapiens (human)		
				PN	JP 199318472-A/2		
				PD	24-NOV-1999		
				PE	22-JAN-1999 JP 1999014949		
				PR	13-MAY-1998 GB 9810373.2,18-AUG-1998 GB 9818009:4 PI		
				PC	CHRISTOPHER D SAZAN, NICOLA BAGESU		
				PC	C12N15/09,A61K31/00,A61K31/00,A61K31/00,A61K31/00,A61K31/70,		
				PC	A61K38/00,		
				PC	A61K39/395,A61K45/00,A61K48/00,C07K14/47,C12P21/02,C12Q1/68,		
				PC	G01N33/15,		
				PC	G01N33/50,G01N33/50,G01N33/566,G01N33/68//C12P21/02,		
				PC	C12R1/91, PC C12N15/00,		

QY 631 LysAlaTyrGluTTPAsnAspAsnGluMetTyrLeuPheArgSerSerValAlaTyrAla 650
 Db 1774 -----GAAGCTCTCATCTCTTGATGTGCTCTGCGCA 1809
 QY 651 MetArgGlnTyrPheLeuLysValLysAsnGluMetLeuPheGlyGluAspVal 670
 Db 1810 CAAGTGAAGATGTTGTT-----TTGTTCTCTACAGGAGAGATGTG 1854
 QY 671 ArgValAlaAsnLeuLysProArgGlySerPheAsnPhePheValThrAlaProLysAsn 690
 Db 1855 CGAGTGGCTAAATTTGAACCAAGATCTCTTAATTTCTTGTCACGACCTAAAT 1914
 QY 691 ValSerAspIleIleProArgThrGluValGluLysAlaIleArgMetSerArgSerArg 710
 Db 1915 GTGCTGATATCATCTCCAGACCTGAAGTGAAGGCGCATCAGGATGTCCGGGCGCT 1974
 QY 711 IleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLeuGlyIleGlnProThr 730
 Db 1975 ATCATATGATGCTTTCGCTCTGATACACACAGCTAGATTTCTGGGATACAGCCACA 2034
 QY 731 LeuGlyProProAsnGlnProProValSerIleTyrPheLeuIleValPheGlyValMet 750
 Db 2035 CTGGAGACCTCTTAACACAGCCCTCTTCCATATGCTGATGTTTGGAGTTGTGATG 2094
 QY 751 GlyValIleValIleValGlyTyrLeuIleLeuIlePheThrGlyIleArgAspArgLysLys 770
 Db 2095 GGAGTGATATGTTGGTGGCTGCTCATCTCGATCTTCTGAGATCAGAGATCGAAGAG 2154
 QY 771 LysAsnLysAlaArgSerGlyLysAsnProTyrAlaSerIleAspIleSerLysGlyLys 790
 Db 2155 AAAAATAAGCAAGAGATGAGAAATCCTTATGCTCCATGATATAGCAAGAGAA 2214
 QY 791 AsnAsnProGlyPheGlnAsnThrAspAspValGlnTyrSerPhe 805
 Db 2215 AATAATCCAGATCCAAAAACATGATGTTCCAGACCTCTTT 2259

RESULT 16
 AX047762 2638 bp DNA linear PAT 15-DEC-2000
 LOCUS AX047762 5 from Patent WO070032.
 DEFINITION Sequence AX047762
 ACCESSION AX047762
 VERSION AX047762.1 GI:11876768
 KEYWORDS
 SOURCE Mus sp.
 ORGANISM Mus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2638)
 AUTHORS Piddington, C.S., Petrie, C.R., Shoemaker, K.E. and Bishop, P.D.
 TITLE Zacc2: a human metalloenzyme
 JOURNAL Patent: WO 0070032-A 5 23-NOV-2000;
 ZymoGenetics, Inc. (US)
 FEATURES
 source location/Qualifiers
 1..2638
 /organism="Mus sp."
 /db_xref="taxon:10095"
 106..2523
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC18992.1"
 /db_xref="GI:11876769"
 /translation="MSSSSVLLSLVAVTAAQSLTEENAKTFLLNNFQEADELSSYSS
 LASVNNITENNOKSSSEAAKMSAYEESQKASQESLOEIQTPILKQLOALOO
 GSSSLASDKNKQNTLITMTSTYSGKVCNPKKPOCILLDEGLDIEIMATSDYNS
 RLMAEGRARVYKGLRPLREYVYLKEMARANNYNDGYGWMGDYABEADQVNT
 RNQLEDEVERTEAEIKPLREHLHVRKKLDYTPSYISPGCLPAHLGDMGRFNT
 NLPLTYFAQCPNIDVDAMNQMGMDEKFEQAEKFEVSGLEPHMTOGFWAMSMIL
 EPADGRKYCEPTAMDLDGDFRIKMTKVTMDNFIHHEMGHIOYDMAYARPPFL
 RGNAGEFHEAVGEIEMSLAATPKIKSIGLPSDFDSEIEINELIKOALITVGLT
 PPTVMLEKRMVPRGRTPKROMMKMKMEKREIYGVVEPLPHDETCDPASLFHVS
 DYSFTIRITRTIYQFOEDALCOAKYNGSLHKCDINSTAGOKLMLSLGNSERH
 KALENVGARMVDKPLINTFQPLEDWLKEQNRSEFGWNTWSPADQSIKRVISL
 KSLAGANAYEWTNNEMPLFRSSVAYAMKRYSIKNTQVPLLEDVIRVSDLKPRVSEY

BASE COUNT 801 a 557 c 611 g 669 t
 ORIGIN
 FPFVSQNVSDYIPRSEVEDATIRMSGRGRINDVFGNDNSLEFIQHPITLEPPYOPPV
 IKLIIIFGVMAIYVVGIIITLIYTGKIGKRRKKREPNREPNDSMDIGGSESNAGRQNS
 DDAQTSF*

Alignment Scores:
 Pred. No.: 1,08e-291 Length: 2638
 Score: 3579.00 Matches: 661
 Percent Similarity: 89.57% Conservative: 60
 Best Local Similarity: 82.11% Mismatches: 84
 Query Match: 83.41% Indels: 0
 DB: Gaps: 0

US-09-978-385-2 (1-805) x AX047762 (1-2638)

QY 1 MetSerSerSerSerTyrPheLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
 Db 106 ATGTCCAGCTCTCTCTGCTCTCTCTCAGCCTTGTTGCTGCTACTGCTCAGTCCCTC 165
 QY 21 IleGluGluGlnAlaLysThrPheLeuAspLysPheAsnIleGluLysPheLeuPhe 40
 Db 166 ACCGAGGAAATGCGAAGACATTTTAAACAACCTTAAATCAGAGACCTGAAGACCTGTCT 225
 QY 41 TyrGlnSerSerLeuAlaSerTyrPasnTyrAsnThrAsnIleThrGluGluAsnValGln 60
 Db 226 TATCAAACTTCACCTGCTCTCTTGAATTAATACACTTACTGAGAAATATGCCCAA 285
 QY 61 AsnMetAsnAsnAlaGlyAspLysTyrPserAlaPheLeuLysGluGlnSerThrLeuAla 80
 Db 286 AAGATGAGTGAAGGTGAGCGACCAATGCTGCTGCTTATATCAAAACAGCTTAAGACTGCC 345
 QY 81 GlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
 Db 346 CAAGCTTTCCTACTACAAAGAAATCCAGACTCCATATCAAGGCTCACTACAGGCCCTT 405
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
 Db 406 CAGCAAAAGTGGGTCTTCAGACCTCTCAGACACAAACAAAGATGGAACAAATTCG 465
 QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
 Db 466 AACACCATGAGCACCATTTACAGTACGAAAGTTTGCAACCAAGAACCCACAAAGAA 525
 QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
 Db 526 TGCTATTACTTGAGCCAGAGATTTGATGAATATATGCGACAGACACTCAACTC 585
 QY 161 ArgLeuTyrPalatrgGluSerTyrPargSerGluValGlyLysGlnLeuArgProLeuTyr 180
 Db 586 AGGCTCTGGGCATGGAGGGGCTGAGGCTGAGAGTTGGCAAGCAGCTGAGAGCGTGTAT 645
 QY 181 GlnGluTyrValIleLeuLysAsnGluMetAlaArgAlaAsnIleThrGluAspTyrGly 200
 Db 646 GAAAGATATGTGTCTGAAACAGATGGCAGACAAACAAATTTTAACGACTATGGG 705
 QY 201 AspTyrTyrPargGlyAspTyrGluValAlaAsnGlyValAspLysTyrAspTyrSerArgGly 220
 Db 706 GATTATTTGAGAGGGGACATGAGACAGAGGACAGACATGCTATACACCTAAC 765
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240
 Db 766 CAGTTGATTTGAAGTGTGTAAGACGTACCTTGCCAGAGATCAAGCCATGTGTGAGCACTT 825
 QY 241 HisAlaTyrValAlaArgAlaLysLeuMetAsnAlaThrProSerTyrIleSerProIleGly 260
 Db 826 CATGCTATGTGAGAGAGAGATGATGATACCTACCTTCTTCAATCAGCCACCTGGA 885
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTyrPylArgPheThrPheAsnLeuTyrSer 280
 Db 886 TGCCCTCCCTGCCCATTTGCTTGATATATGAGGGGTGATTTTGGCAAAATCTGTACCT 945
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300

```

|||||
Db 946 TTGCTCTTCCTTCGCACAGAAACCAACATAGATGTACTGATGCAATGATGAAATCAG 1005
QY 301 Alaripaspalaginaragiliepbelysglualagilysphepvalservalglyleu 320
Db 1006 GCGTGGGATGAGAAAGGATTTTTCAGAGAGAAATCTTTTCTGTTGGCCCTT 1065
QY 321 Proasmetrhinginslypethrpgluasensermetleuthraspprogiluasnavalgin 340
Db 1066 CCTCATATGACTCAAGATTTCTGGCAAACTCTATGCTGATGATGACGACGAGATGGCCGG 1125
QY 341 Lysalavalalyspshspprothralatrpaspheuglylyscllyaspheargileumet 360
Db 1126 AAGATTGCTGCGACCCACAGCTTGGGATCTGGGACACGAGACTTCAGATCAATCAAGATG 1185
QY 361 CysthrlysalThrmeltaaspapheleuthralahisglumetglshislegin 380
Db 1186 TGTACAAAGSTGCACATGGACAACTTCTTGACAGCCATCAGAGATGGACATCCAA 1245
QY 381 TyrAspmetalaTyralaalaginProphleuleuArgasngllyalaasnuglylpe 400
Db 1246 TATGACATGSCATATGCGACGCAACTTCTCTCTAGAAACGAGACCATGAGGCTTC 1305
QY 401 HisglualavalgllylulemetSerleuSerAlaIatThProlyshisLeuysSer 420
Db 1306 CATGAAGCTGTGGAGAAATCATGTCATTTGTCAGACTACCCCAAGCATCTGAATCC 1365
QY 421 IleglyleuleuSerProasppheingluasasnugluthrGluliasnPhleuleu 440
Db 1366 ATTGCTCTTCTGCCATCCATTTTTCAGAGATGCGGAAACAGAGATTAACCTCTACAGT 1425
QY 441 LysglualaleuthrIlevalgllyhrleuProphethrThrmetyrmetleugluysrparg 460
Db 1426 AAACAGCATTTGACATTTGTGAAACACTACCGTTTACTTACATGTTAAGAGAGTGGAGG 1485
QY 461 TrpmetvalPheylscllylulepPolysaspglIntpmetlylsystrpplumet 480
Db 1486 TGGATGCTTCTGCGGGGTAAATTCACAAAGAGCAGTGGATGAAAAAGGTGGGAGATG 1545
QY 481 LysarglulilevalgllyalvalgluprovalProhisaspgluthrlyrcysasppro 500
Db 1546 AAGCGGAGATGTTGGTGTGTGTGGAGCTTCTGCTCATGATGAACATACGTGATCCCT 1605
QY 501 AlaserleuphehisvalserasnsaprySerpheilearglyrThrarghrleu 520
Db 1606 GCATCTCTGTCATGTTCTATGATTAATCACTCATTCATTCATTAACAAGAGCATT 1665
QY 521 TyrGlnpheglnPheglnglualeucysglnalaialylshisgluglyProleuhis 540
Db 1666 TACCATTCACATTCACAGAACGCTTTGTCAGAGCAGCTAAGTATATGTTCTCTGCAC 1725
QY 541 LysCysaspiliserAsnserserThrgluagilysleuPheasmetleuargleu 560
Db 1726 AATATGACATCTCAATTTCCACTGACCTGGGACGAAAGTTGCTCAAGATGCTGCTCTT 1785
QY 561 GlylysSerGluProTrpThrleuAlaleugluasnavalvalgllyalylasmetasn 580
Db 1786 GGAATTTAGAGCCCTGAGACCAAGCTTGGAAATGTGTAGAGCAAGCAATATGTGAT 1845
QY 581 ValArgProleuleuAsnThrPheglubProleuphemTrpPleuylsAspGlnasulys 600
Db 1846 GTAAACACACCTGCTCAATTAATCTCCAAACGCTTGTGACTGGCGAAGACCAACAGCA 1905
QY 601 AsnsersphevalgllyrSerThrasptRserProlyrAlaaspglnSerIleuysVal 620
Db 1906 AATCTTTTGGGGGTGAGAACACTGAATGAGGCCATATGCCGACCAAGCATTAAGATG 1965
QY 621 ArgIleSerleuysSerAlaleuglyAspysalatyrglurtPasnaspasnuglmet 640
Db 1966 AGGATTAACCTTAATATGAGCTTGTGAGCTTAATGATGATGAGCAACAGCAATG 2025
QY 641 TyrleupheargserSerValAlaTyralametrGlnTrpPheleuylsVallyasn 660
|||||

```

```

Db 2026 TTCCTGTTCCGATCATCTGTTCCATATGCCATAGAGAAAGTATTTTCAATATCAAAAAC 2085
QY 661 GlnmetIleleuphegllygluliaspvalArgvalAlaAsnleuysProargIleser 680
Db 2086 CAGACAGTTCCTTTCTTAAGAGAGATGTAGCAGAGACGATTTCAACCAAGATCTCC 2145
QY 681 PheAsnphesphevalThralarPolysasnavalseraspilIleleproargThrgluval 700
Db 2146 TTCTACTTCTTTTGTCACTCACCCACCAAAATGTCTGTGATGATCATCTTAAGAAAGTAA 2205
QY 701 GlulysAlaIleargmetSerargserArgileasnaspalapheargleuasnaspasn 720
Db 2206 GAAAGTGCATAGAGATGCTGCGGCGCCCATCATGATGCTTTGGCTTAATGATTAAC 2265
QY 721 SerleugluPheleuglylIlegluProThrleuglylProasnglnProProvalSer 740
Db 2266 AGCCTGAGATTTCTGGGATTCACCCAAACACTTGAGCCACTTACCACCTCCGTGTCACC 2325
QY 741 IletrpleuIlevalPhegllyalvalmetgllyalIlevalvalgllyIlevalIleleu 760
Db 2326 ATATGCTGATTAATTTTGTGTGTGTGTGATGGCACGTGATGTTGGCATCATCATCTG 2385
QY 761 IlephethrglyleargAsparqlslyslslysnalysAlaargserglygluasnPro 780
Db 2386 ATTGCTACTGGGATCAAAAGTCTGCAAGAGAAATGAAACAAAGAGAGAACACCT 2445
QY 781 TyralaSerIleaspiliserlysglygluasnasnProgllypheglnAsnThrasp 800
Db 2446 TATGACTCGATGAGCATTTGAGAAAGAGAGAAACATGAGATTCACAAACAGATGAT 2505
QY 801 ValGlnThrserPhe 805
Db 2506 GCTCAGACTTCTCTT 2520

RESULT 17
BC026801 2739 bp mRNA linear ROD 07-ANG-2002
LOCUS BC026801
DEFINITION Mus musculus, RIKEN cDNA 2010305L05 gene, clone MGC:25940
IMAGE:4236529, mRNA, complete cds.
ACCESSION BC026801
VERSION BC026801.1 GI:20071163
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryote: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLES 1 (bases 1 to 2739)
JOURNAL Direct Submission
Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/BLM at: http://image.llnl.gov
Series: IRAX Plate: 31 Row: e Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF

```

FEATURES analysis, similarity but not identity to protein.
 source Location/Qualifiers
 1. 2739

/organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:25940 IMAGE:4236529"
 /tissue.type="Kidney, normal, 5 month old male mouse."
 /clone_lib="NCI_CGAP_Kid14"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 176. 2593

CDS

/codon_start=1
 /product="Riken cDNA 2010305L05 gene"
 /protein_id="AAH26801.1"
 /db_xref="gi:20071164"
 /db_xref="locusid:70008"
 /translation="MSSSSWLLSLVAVTTAQSJTEENAKTFLENNQAEADISTOSS
 LASNYNTNITEENAKMSEAAKWSAFYEQOSTAQSFSIQETQPIKROQLALQO
 SSSALSDAKKCOLNTILMTSTISTGKYNPNPNPOECLEBGLDEIMASTDYNS
 RLMAEGRAEVGKOLRPLYEYVVLKNEKMARANNNDYGDYEAEGADGYNR
 RNQIIEDEPFAEIKPLYEHLHAYVRKLMIDTYSYISPGCLPAHLGDMWMSMT
 NLYPLVFAKPNIDYTDAMNMGDAERFEQAEKPFYVGLPHYTGFWMSMT
 EPADGRKYCHPTAMDGHGDFRIKCTKYMDNPLRHHMHGHIQYDMAARPFL
 RENGAGEHEAVGETMSLAATPKHLSTGLPDSFQEDSETLEFLKQALITVGL
 PPTVLAERKMWVREGELPKQMKKWKEMKREIVVEPLHDEYICDPSLEHVN
 DYSFIRYRTIYQFQALCOAKYKGLSKDINSSTAGOKLKMSLGSEBW
 TKALENNVGARMVDYKPLNFOPFLDKLQNNNSNTGEMSPADDSIKYRIL
 KSALENNVGARMVDYKPLNFOPFLDKLQNNNSNTGEMSPADDSIKYRIL
 EFVTSQVNSVIRSEVEDAIRMSRGINDVPLNLSLEFLGHPLEPPQPPVT
 IWLIFGVVMAVVVGIITLVITGKGRKKNETKRENPIDSMIDIGKESNAQFONS
 DDAQTSF"

BASE COUNT 832 a 577 c 620 g 710 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,13e-291 Length: 2739
 Score: 3579.00 Matches: 661
 Percent Similarity: 89.57% Conservative: 60
 Best Local Similarity: 82.11% Mismatches: 84
 Query Match: 83.41% Indels: 0
 Gaps: 0

US-09-978-385-2 (1-805) x BC026801 (1-2739)

QY 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
 Db 176 ATGTCAGCTCCCTCGCTCCCTTCACAGCTTGCTGCTACTACCTCCTCAGTCCCTC 235
 QY 21 TleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGlnLysPhe 40
 Db 236 ACCGAGAGAAATGCCAAGACATTTTAAACAACCTTAAACAGAGAGCTGAAGACCTGCT 295
 QY 41 TyrGlnSerSerLeuLeuSerTrpAsnLysThrAsnLysThrGlnGluGlnValGln 60
 Db 296 TATCAAAAGTTCCTCTCTGCAATTTATTAATACTAACAATTAAGTGAAGAAATGCCCA 355
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAla 80
 Db 356 AAGATGATGAGCTGCAGCAAAATGCTGCTCTTTATGAAGAACAAGCTGAAGACTGCC 415
 QY 81 GlnMetLysTrpLeuGlnGlnLysGlnGlnSerValLysGlnGlnLeuGlnAlaLeu 100
 Db 416 CAAAGTTCCTCCTACCAAGAAATCCAGACTCCGATCATCAAGCTCAACCTAAGAGCCCT 475
 QY 101 GlnGlnAlaGlnLysSerValLeuSerGlnAspLysSerLysArgLeuSerThrLeu 120
 Db 476 CACCAAGAGTGGCTTCAGCACTCAGCAGACAGAAACAAGAGTTGACCAATTCG 535
 QY 121 AsnThrMetSerThrLysThrGlnLysValLysAsnProAspAsnProGlnGln 140
 Db 536 AACACCATGAGCACTTAACTACTGAGAAAGTTTGCACCCAAAGAACCCACAGAGA 595

QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGlnLysMetAlaAsnSerLeuAspTrpAsnGln 160
 Db 596 TGCCTTATTAATTAATGAGCCAGATTTGATGAATATATGGCCAGACAGACACTCACTC 655
 QY 161 ArgLeuTrpAlaTrpGlnSerTrpArgSerGlnValGlyLysGlnLeuArgProLeuLys 180
 Db 656 AGGCTCGGGCATGGGAGGGCTGGAGGCTGAGAGTGGCAAGACGATGAGGCGGTGTAT 715
 QY 181 GlnGluTrpValValLeuLysAsnGlnMetAlaArgAlaAsnHisThrGlnLysGln 200
 Db 716 GAGAGATATGCTCTCCGAAAACGAGATGGCAAGACCAACATTAATCAAGCTATGG 775
 QY 201 AspTrpTrpArgLysAspTrpGlnValAsnGlyValAspGlyLysTrpSerArgLys 220
 Db 776 GATATATGAGAGAGGAGCATATGAAGCAGAGGAGCAGATGCTAACAATTAACCTTAAC 835
 QY 221 GlnLeuLysGlnAspValGlnHisThrPheGlnGlnLysProLeuLysGlnHisLeu 240
 Db 836 CAGTTGATGAAGTATGAGAACGATCTTCGACAGATCAAGCATATGATGACATCTT 895
 QY 241 HisAlaTrpValArgAlaLysLeuMetAsnAlaLysProSerTrpLysSerProLysGly 260
 Db 896 CATGCTATGAGAGAGAGGAGATGATGATGATACCTCCTTACATACACCCACTGGA 955
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuLys 280
 Db 956 TGCTCCCTGCCCTTGGCTTGGATATGAGGATGATTTGGCAAAATCTGTACCT 1015
 QY 281 LeuThrValProPheGlyGlnLysProAsnLysAspValThrAspAlaMetValAspGln 300
 Db 1016 TTGACTGTCCCTTGGACAGAAACCAACATATGATTTACTGATGAAATGATGATCAG 1075
 QY 301 AlaTrpAspAlaGlnArgLysPheLysGlnAlaGlnLysPhePheValSerValGlyLeu 320
 Db 1076 GCGTGGATGCAAGAAAGATATTCAGAGAGCAGAGAAATCTTGTGTTCTGTGGCCTT 1135
 QY 321 ProAsnMetTrpGlnGlyPheTrpGlnLysSerMetLeuThrAspProGlnValGln 340
 Db 1136 CCTCATATGATCTCAGAGATCTGGGCAACCTTATGCTGATGAGCCAGCAGATGGCCGG 1195
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlnAspPheArgLysLeuMet 360
 Db 1196 AAGTGTGTCTCCACCCACAGCTTGGAGTCTGGAGACAGAGACCTCAGAAATCAAGATG 1255
 QY 361 CysThrLysValThrMetAspPheLeuThrAlaHisGlnMetGlyHisLysGln 380
 Db 1256 TGTCAAAAGTGCACAAATGACAACTCTTGACACCCATCAGCAGAGTGGAGACATCCAA 1315
 QY 381 TyrAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPhe 400
 Db 1316 TATGACATGCAATATGCCAGCAACCTTCTGCTAAGAAACGAGCAATGAAGGCTTC 1375
 QY 401 HisGlnAlaValGlyLysLysMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 Db 1376 CATGAAGCTGTGGAGAAATCATGCTTCTGCTACCTACCCCAAGCATTCGAATTC 1435
 QY 421 IleGlyLeuLeuSerProAspPheGlnGlnAspAsnGlnLysGlnLysPheLeuLeu 440
 Db 1436 ATTGCTCTTCCGATCGATGATTTCAAGAAATGCAAGACAGAGATTAACCTTCATCTG 1495
 QY 441 LysGlnAlaLeuThrLysAlaGlyThrLeuProPheThrLysMetLeuGlnLysTrpArg 460
 Db 1496 AAACAGGATTCGACATATGTTGGAAACACTTCTTACTTAACATGTAAGAGAGTGAAG 1555
 QY 461 TrpMetValPheLysGlyLysLysLeuLysAspGlnLysTrpMetLysLysTrpPheMet 480
 Db 1556 TGGATGCTCTTCCGGGGTGAATTCCTCAAGAGCAGAGTGAAGAAAGTGGTGAAGATG 1615
 QY 481 LysArgGlnLysLysAlaGlyValAlaGlnProValProHisAspGlnLysTrpCysAspPro 500
 Db 1616 AAGCGGAGATCGTGGTGTGTGGAGGCTTCCCTATGATGATGAACATCTGAGACCT 1675
 QY 501 AlaSerLeuPheHisValSerAsnAspLysSerPheLeuArgLysTrpThrArgThrLeu 520

Db 2587 TGCTCAGACTTCCTT 2602

RESULT 19
AX047765 2638 bp DNA linear PAT 15-Dec-2000
LOCUS AX047765
DEFINITION Sequence 8 from Patent WO0070032.
ACCESSION AX047765
VERSION AX047765.1 GI:11876771
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2638)
AUTHORS Piddington, C.S., Perrie, C.R., Shoemaker, K.E. and Bishop, P.D.
TITLE Zace2: a human metalloenzyme
JOURNAL Patent: WO 0070032-A 8 23-NOV-2000;
Zymogenetics, Inc. (US)

FEATURES
source location/Qualifiers
1..2638
/organism="Mus sp."
/db_xref="taxon:10095"
106..2523
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC18993.1"
/db_xref="GI:11876772"
/translation="MSSSMILLIVAVTTAOSTEENAKTFINNTNOEADISSSS
LASNNYNTITEENAAKMSAFYEQSKTOSFESLOETPTPIKROLQALQO
SGSSALADKNKOLNTLNTMSTISTGKVCANRNPOECLELPEGLDEIMASTDYS
RUMAEGRRAVAGKQLPLEEVVILKNEAARNANNNDYGRGDEAGADGYNN
NQIIEVEREFAIKPLVLEHYAVKRLMDYPSYISTGCLPAHLGDMGRMT
EPAGRRVYVCPAMDIGHDFRIKCTKVYMDNPFILAHHEMGIHIOYDAVYRQPLT
RNGANECFHEAVGEMISLAAATPKHLKSGILPSDPODSEREINLLQATITGTL
PFTMLEKMMWVEGEIPEKQMKKMKKPRIVGVPEPLPRDEYCDPASLFEVSN
DYSPITRYTFOFOFOALCOAKYNSLRKCDISNTEAGQKLKMLSGNSEPW
TEALENVVGRNMDYKPLNFPPLDMLEKQNRNSFVWNTWESPDAOQSIRVSI
KSLGANAAYRWNTNMEFLFRSVAVMKRYSSIKNQVPLEEDYRVSDELPRVSY
PFVTSPOVSDVPRSEVEDAIRMGRINDYFGINDLEFLGIHPTLEPYOPPT
TWLIRGVVALVVGIIILITVIGIRKKKNETKREPNYSMDIGKESNAQRONS
DDAQTSS"

BASE COUNT 797 a 557 c 615 g 669 t
ORIGIN

Alignment Scores:
Pred. No.: 3.55e-290 Length: 2638
Score: 3561.00 Matches: 659
Percent Similarity: 89.32% Conservative: 60
Best Local Similarity: 81.86% Mismatches: 86
Query Match: 82.99% Indels: 0
Gaps: 0

US-09-978-385-2 (1-805) x AX047765 (1-2638)

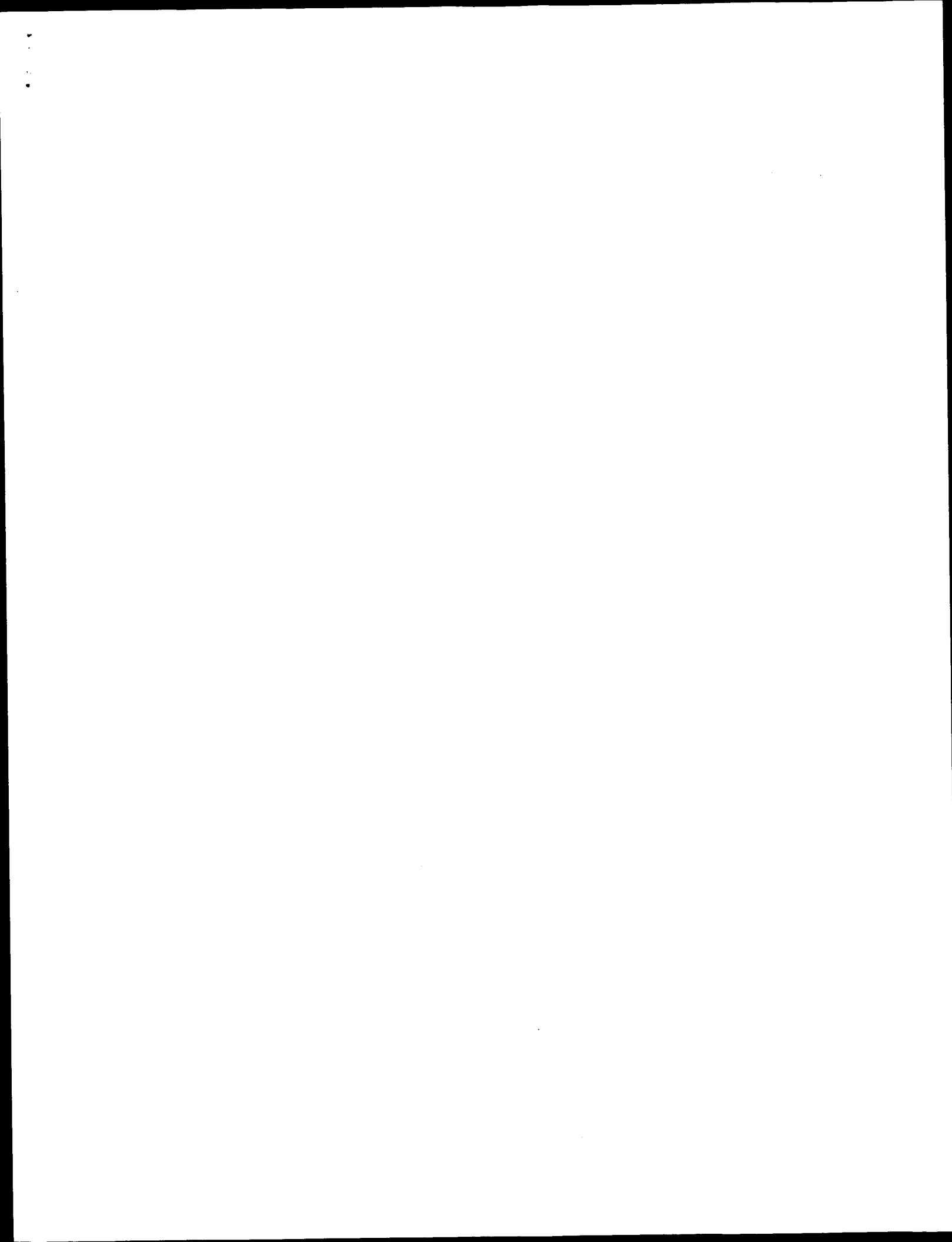
QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
Db 106 ATGTCAGCTCCTCCCTGCTTCACGCTTGCTGTTACTACTCTCAGTCCCTC 165
QY 21 IlleGlnGlnAlaLysThrPheLeuAspLysPheAsnHisGlnAlaGlnAspLeuPhe 40
Db 166 ACCAGGAAATGCCAAGACATTTTAAACACTTTAAACAGAGGCTGAGACCTGCT 225
QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGlnGlnLysValGln 60
Db 226 TATCAAGTTCACCTGCTGTTTGGAAATTTATTAATACTTAACATTACTGAAATAATCCCAA 285
QY 61 AsnMetLysAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAla 80
Db 286 AAGATGAGTGGCTGACGACCAATATGCTGCTTTTATGAGAAACAGCTCAAGCTGCC 345
QY 81 GlnMetTyrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
||| ::: ||||||||||||||| :||| |||||||||||

Db 346 CAAAGTTTCCTACACAGAAATCCAGACTCCGATCATCAACCTGCAACGCTT 405
QY 101 GlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThrIleLeu 120
Db 406 CACCAAGAGTGGGCTTCACACACTCTCCAGCAGCAAGAACAAACAGTCAACCAATCTG 465
QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnProGlnGlu 140
Db 466 AACACCATGAGCCACCTTTTACAGTACTGGAAAAGTTTCCAAACCCAAAGAACCCACAGAA 525
QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGlu 160
Db 526 TGCTTATTACTTACGACGACAGATTTGATGAATAATGGGACAGACAGCATACAACTCT 585
QY 161 ArgLeuThrAlaThrGlnSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyr 180
Db 586 AGGCTTGGGCAATGGAGGGCTGGAGGGCTGAGTTGGCAAGCAGCTGAGGCCCTGTAT 645
QY 181 GlnGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAspTyrGly 200
Db 646 GAAGAGTATGTGCTCTGCTGAAAAACGAGATGGCAAGACAAACATTTAAGACTATGGG 705
QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
Db 706 GATTATTGGAGAGGGGAGCTATGAACAGAGGAGCAGATGGCTCAACTATTAACCTTAAC 765
QY 221 GlnLeuIleGluAspValAlaGlnIleThrPheGlnGlnIleLysProLeuTyrGluHisLeu 240
Db 766 CAGTTGATGATGAAGATGTGAACGATACCTCCCAAGATCAACCAATGTATGAGCATCTT 825
QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db 826 CATTGCTTGTGAGAGGAGAACTGATGATACCTACCTCTCCATCAACACCCACCTAC 885
QY 261 CysLeuProAlaHisLeuLeuGlnAspMetTrpGlyArgPheThrPheAsnLeuTyrSer 280
Db 886 TGCTCCCTCCGCAATTTGCTTGATATGATGGGGTATGATTTGGACAAATCTGTACCT 945
QY 281 LeuThrValProPheGlnGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 946 TTGACTGTTCCTTGGCAGCAAAACCAACATGATGTTACTGATCAATGATGATGATG 1005
QY 301 AlaTrpAspAlaArgIlePheLysGlnAlaGlnLysPhePheValSerValGlyLeu 320
Db 1006 GGCTGGATGCGAAGAAAGATTTTCAAGAGGCGAGAAATTTCTTCTGTTGCTGCTT 1065
QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyLysValGln 340
Db 1066 CCTCATATGACTCAAGGATTTCTGGCAAACTCTATGCTGACTAGCCACACAGATGGCCGG 1125
QY 341 LysAlaValCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1126 AAAGTTGCTGCCACCCCAAGCTTGAGATCTGGACAGCGAGACTTCGAATCAAGATG 1185
QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnIleMetLysHisLeuGln 380
Db 1186 TGTCAAAGGTTCACAAATGAGCAACTCTTGACAGCCCATCAGAGATGGGACATCTCAA 1245
QY 381 TyrAspMetValTyrAlaAlaGlnProPheLeuLeuArgAsnGlyValAsnGlnLysPhe 400
Db 1246 TATGACATGGCATATGACAGCAACCTTCTGCTGAAGAACGAGGACCAATGAAGGTTTC 1305
QY 401 HisGlnAlaValAlaGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
Db 1306 CATGAAGCTTGAGAAATCATGTCATTTCTGACAGCTACCCCAACAGCATGTGAATATCC 1365
QY 421 IleGlyLeuLeuSerProAspPheGlnGlnLysAsnGlnTyrGlnIleAsnPheLeuLeu 440
Db 1366 ATGGAGCTTCTGCAACCATTTTCAAGAAATAGCAAAACAGAGATAAATCTCTACTG 1425
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTyrArg 460
Db 1426 AACACGCAATGACATTTGTTGAAACACTACCGTTTACTTACATGTTTGAAGAACTGGAGG 1485

Db 541 GARGARTAYCTNGTNTYNNARAAYGARATGGCGMGGCAATCATTAAGARATTAATGCGN 600
 Qy 201 ASPTTYTTPARGGLYASPTTYGLUVALASNGLYVALASPGLYTYRASPTRYSEARGLY 220
 Db 601 GATTAATGGMGNGNGAAYRAYGARGNAYGGNGTNGAYGGNTAYGAYTAYWSNMNGCN 660
 Qy 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleValSerProLeuThrGluHisLeu 240
 Db 661 CARYNATHTGARGAYCTNGARCAACNTTGGARARATHAACCTTATTAAGARCAAYTN 720
 Qy 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
 Db 721 CAAGCAATAGTNGMGCNNAARTNATGAAGCNTAYCCNWSNTATATWSMCCNATHGCGN 780
 Qy 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpHisAsnLeuTyrSer 280
 Db 781 TGYTTCGCCGNCATYTTNTTNGNGATATGTGGGNGMGNNTTGTGACNAAAYTTTAYWSN 840
 Qy 281 LeuThrValProPheGlyGluLysProAsnIleAspValThrAspAlaMetValAspGln 300
 Db 841 YTNACGTCNCCNTTGGNCAARARCCNNAAYATHGATGNACNGAYGCAATGCTNGATCAR 900
 Qy 301 AlaTTPAspAlaGluArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320
 Db 901 GCGTGGGATGCGNCAAMGNATHTTAAARGARCGNARARATTTTGTWMSNGTNGCYTN 960
 Qy 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
 Db 961 CCNAAAYATGACNCARGNTTGTGGARAAYSNATGTATMCNAGAYCCNCGNAAAGTICAR 1020
 Qy 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
 Db 1021 AARGNGTNTGCAACNACGTCGGAAYTNGNNAARGNGAYTTTGMNATHTYTATG 1080
 Qy 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHisIleGln 380
 Db 1081 TGYACNNAARGTNACNATGATGATTTTATNACNCCNCAATCAYGARATGGGNCAYATHCAR 1140
 Qy 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
 Db 1141 TAYGAYATGGCNTAYGCGNCNCAACNCCNTTATNTNMGNAYGGNGCNAAAYARGCNTTY 1200
 Qy 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 Db 1201 CAYARGCNGTNGNGARATHTATGWSNTTWSMCCNCCNCAARCAAYTNAARWSN 1260
 Qy 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440
 Db 1261 ATHGNGTNTTWSNCCNCAATTTTCARGARATTAAGARCCNARATHTAATTTTNTNTN 1320
 Qy 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460
 Db 1321 AARCARCNGTNNACNATHGNGNACNNTTTCNCAATATATGTYNGARARARGTGMGN 1380
 Qy 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrpGluMet 480
 Db 1381 TGGATGCTNTTAYARGNGNGARATHCCNAAARGATCATGATGAARARARGTGGGARRAG 1440
 Qy 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500
 Db 1441 AARGNGARATHGNGNGTNGARCCNCGTNCNCAATGARGARCCNATAYTGAYCCN 1500
 Qy 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520
 Db 1501 GCNWSNTNTTTCAYGNGNWSNAAAYGAYTAYWSNTTATATHTMGNTATATAYACNMGNACNTN 1560
 Qy 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis 540
 Db 1561 TAYCARTTYCARTTYCARGARGCNGTNTTGTGCARGCNGCNAAARCAATGARGCNGTNTTCAY 1620
 Qy 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560

Db 1621 AARTGCAVATHTWSNAAYSNACNGARCGCGNCARARATNTTAAATGTYNMGNYN 1680
 Qy 561 GlyLysSerGluProTrpThrLeuAlaGluAsnValValGlyAlaLysAsnMetAsn 580
 Db 1681 GGNAAKWSNCAACNCTGGACNCTTNGCNTTNGARRAAYGNTGNGCNAAARAAATGAAAY 1740
 Qy 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600
 Db 1741 GTNMGNCNTYNTTNAAYATATYTGARCCNTNTTACNTGGTYNNAARGATCAARAAAR 1800
 Qy 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620
 Db 1801 AATWSNTTGTGNGTNGMNSNACNGAYTGMSNCCNTAYCGNGAYCARWSNATHAARGTN 1860
 Qy 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet 640
 Db 1861 MGNATHTWSNTTAAARWSNCCNTTNGCNATAYARCGNTAYGARTRGAAAYGATAYGARATG 1920
 Qy 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660
 Db 1921 TAYTNTTGTGMSNWSMNSNGTNGCNATYGCNATGMCNCAATYTTTNTAARGTNARAAY 1980
 Qy 661 GlnMetIleLeuPheGlyGluLysAspValArgValAlaAsnLeuLysProArgIleSer 680
 Db 1981 CARATGATHTTNTTGTGNGARGARATYCTNMGNGTNGCNAAAYTTTNAARCCNMGNATHWSN 2040
 Qy 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleLeuProArgTrpGluVal 700
 Db 2041 TTYAAYTTTGTGTACNCGCCNCAARAAAYGTNMSNCAATHTATHCNMGNACNGARGTN 2100
 Qy 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
 Db 2101 GARARCGNATHMGNATGMSNWSMNGNATHAAYGAYCCTTNGNTYNNAAAYGAYAY 2160
 Qy 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
 Db 2161 WSNYTGARTTYYTNGGNATHCARCNAACNTTNGCNCNCAAYCARCCNCCNCGTNGWSN 2220
 Qy 741 IleTrpLeuIleValPheGlyValAlaMetGlyValIleValValGlyIleValIleLeu 760
 Db 2221 ATHTGCTTNAATHTGNTTGGNGTNGTATGCGNTATHTGNTGNGNATHTGNTATHTYN 2280
 Qy 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnPro 780
 Db 2281 ATHTTACNCGNATHMGNGAYMGNAARAAARAAARAAARCCNMGNWSNCGNGARRAAYCCN 2340
 Qy 781 TyrAlaSerIleAspIleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAsp 800
 Db 2341 TAYGCMNSNATHGAYATHTWSNAAARGNGARAAAYAACCGNGNTTTCARAAAYACNGAYAY 2400
 Qy 801 ValGlnThrSerPhe 805
 Db 2401 GTNCARACNWSNTTY 2415

Search completed: March 7, 2003, 09:08:57
 Job time : 4946 secs



GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 7, 2003, 06:33:34 ; Search time 406 Seconds
(without alignments) 4465.169 Million cell updates/sec

Title: US-09-978-385-2
Perfect score: 4291
Sequence: 1 MSSSSWLLSLVAVTAAGST.....ISKGNPNPFGNTDDVQTSF 805

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus.p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09978385/rnat.28022003.104716.1332/app-query.fasta.1.967
-DB=N.Geneseq.101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -DOOPCL=0
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=20
-MODE=LOCAL -OUTFMT=ptc -NOMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09978385 -GCN 1.1 -36 -ftrnal -28022003.104716.1332 -NCPU=6 -ICPU=3
-NO_XLPPY -NO_MMAB -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4291	100.0	2418	21	AAZ59465	Human MPR015 codi
2	4291	100.0	3334	22	AAC84366	Human Zace2 protei
3	4291	100.0	3396	21	AAAI2764	CDNA encoding a hu
4	4291	100.0	3396	22	AAAD0758	Human angiotensin
5	4291	100.0	3396	24	ABR87623	CDNA encoding huma
6	4291	100.0	3396	24	AAD32586	Human ACE-2 full-1
7	4142	96.5	2920	22	AA521279	Human ACE-2 full-1
8	4061	94.6	2920	22	AA514890	Human CDNA sequenc
9	4013	93.5	2911	22	AA514890	Human CDNA sequenc
10	3740.5	87.2	2262	21	AAZ59465	Human CDNA encodin
11	3579	83.4	2638	22	AAC84368	Human MPR015 codi
12	3561	83.0	2638	22	AAC84368	Human MPR015 codi
13	3509	81.8	2415	22	AAC84370	Mouse Zace2-5 prot
14	3291	76.7	2350	24	ABK30270	Mouse Zace2-10 pro
15	3119	72.7	3474	22	AA542515	Human G-protein-co
16	2899	67.6	2415	22	AA542515	Human CDNA encodin
17	1344	31.3	2477	12	AAQ10338	Human Zace2-5 prot
18	1337	31.2	2477	12	AAQ10338	Human Zace2-5 prot
19	1337	31.2	2477	12	AAQ10338	Human Zace2-5 prot
20	1337	31.2	2477	12	AAQ10338	Human Zace2-5 prot
21	1337	31.2	2477	12	AAQ10338	Human Zace2-5 prot
22	1336	31.1	4020	19	AAV41320	Human angiotensin
23	1334	31.1	3939	22	AA535850	Human angiotensin
24	1334	31.1	3939	22	AA535850	Human angiotensin
25	1275	30.5	3942	20	AA506057	Human angiotensin
26	1275	29.7	5005	22	AAH57430	Human angiotensin
27	1240.5	28.9	2568	24	AAK99739	Human angiotensin
28	1086	25.3	2089	23	ABLI4379	Human angiotensin
29	1057	24.6	2074	16	AA082948	Human angiotensin
30	1028	24.0	2450	22	ABLI4379	Human angiotensin
31	961	22.4	9005	22	ABLI6677	Human angiotensin
32	941.5	21.9	5632	23	AAH77873	Human angiotensin
33	919.5	21.4	5050	23	ABLI4378	Human angiotensin
34	841	19.6	2379	23	ABLI6696	Human angiotensin
35	769	17.9	561	24	ABO57861	Human angiotensin
36	721	16.8	2082	21	AA46692	Human angiotensin
37	721	16.8	2082	22	AA514169	Human angiotensin
38	721	16.8	2082	22	AA514169	Human angiotensin
39	715.5	16.7	1395	22	AAH77873	Human angiotensin
40	597.5	13.9	42573	23	ABO57861	Human angiotensin
41	597.5	13.9	42573	23	ABO57861	Human angiotensin
42	502.5	11.7	2025	23	ABLI6696	Human angiotensin
43	478	11.1	313	20	AAV41320	Human angiotensin
44	476	11.1	1836	23	ABLI27143	Human angiotensin
45	469	10.9	467	19	AAV09277	Human angiotensin

ALIGNMENTS

RESULT 1
AAZ59465
ID AAZ59465 standard; DNA; 2418 BP.
XX
AC AAZ59465;
XX
DT 11-APR-2000 (first entry)
XX
DE Human MPR015 coding sequence #1.
XX
KW MPR015; treatment; hypertension; human; myocardial disease; apoplexy;
KW heart disease; apoplexy; heart disease; nervous denaturation; ds;
KW Alzheimer's disease; hormone; cytokine.
XX
OS Homo sapiens.
XX
PN JF11318472-A.
XX

Db 1621 AATGTCACATCTCAAACTGTACAGAACTGACAGAACTGTTCAATATGCTGAGGCTT 1680
 QY 561 G1YLSerSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetLsn 580
 Db 1681 GGAATAATCAGAACCTTGACCTTACATGATGGAATAATCTTGTAGAGCAAGAACATGAAAT 1740
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600
 Db 1741 GTAAGCCACTGCTCACTACTTGTAGACCTTATTTACTGCTGGAAGACCAAGAACAG 1800
 QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTrpThrLaspGlnSerIleLysVal 620
 Db 1801 AATCTTTTGGGATGAGTACGACTGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpPasnAspAsnGluMet 640
 Db 1861 AGGTAAGCTTAATAAGCTCTTGGAGATTAAGCATATTAATGGAACGACATGAAATG 1920
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgIleTyrPheLeuLysValLysAsn 660
 Db 1921 TACCTGTCGATCATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
 QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680
 Db 1981 CAGATGATCTTTTGGGAGAGAGATGTCGAGTGCCTAATTTGAAACCAACAACTCTCC 2040
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgTrpGluVal 700
 Db 2041 TTTAATTTCTTTGTCACTGACCTTAATAATGTCTGATTCATCTTCTAGAACTGAACTT 2100
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
 Db 2101 GAAAGGCCATCAGATGTCCTCCGAGCCGATATCATGATGCTTCCCTGTAATGACAAAC 2160
 QY 721 SerLeuLupheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
 Db 2161 AGCCTAAGTTCTGGGGAATACACCAACCTTGACCTCTTAACCCAGCCCTGTTTCC 2220
 QY 741 IleTrpLeuIleValPheGlyValIleMetGlyValIleValGlyIleValIleLeu 760
 Db 2221 ATAAGGCTGANTGTTTGGAGTTGTGATGGAGTGTATGTTGTTGTCATTCATCTG 2280
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780
 Db 2281 ATCTTCACCTGGATCAGAGATCCGAAAGAAATTAACCAAGAGTGGAAATCTCT 2340
 QY 781 TyrIleSerIleAspIleSerLysGlyLysAsnProGlyPheGlnAsnThrAspAsp 800
 Db 2341 TATCCTCTCATCGATATTAGCAAGAGAAATATATCCAGATTCCAAAACACTGATGAT 2400
 QY 801 ValGlnThrSerPhe 805
 Db 2401 GTTCAGACCTCTTT 2415
 RESULT 2
 AAC84366
 ID AAC84366 standard; cDNA; 3334 BP.
 XX
 AC AAC84366;
 XX
 DT 19-MAR-2001 (first entry)
 XX
 DE Human Zace2 protein encoding cDNA.
 XX
 Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
 KM zinc metalloproteinase; blood pressure; zinc protease; hypertension;
 KM ventricular systolic dysfunction; renal impairment; heart failure;
 KM scleroderma renal crisis; atherosclerosis; . antiinflammatory; human;
 KM antiarthritic; bradykinin inactivator; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers

FT CDS 35..2452
 FT /*tag= a
 FT /product= "Zace2"
 PN MO200070032-A1.
 XX
 XX 23-NOV-2000.
 XX
 XX 03-MAY-2000; 2000MO-US11932.
 XX
 PR 13-MAY-1999; 99US-0311482.
 PR 27-AUG-1999; 99US-0384706.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
 DR WPI; 2001-025018/03.
 DR P-PSDB; AAB48095.
 XX
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 PT associated with inflammation such as arthritis and enterocolitis -
 XX
 XX Example 1; Page 95-100; 125pp; English.
 CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood
 CC pressure regulation and fertility. Zace2 can be expressed by standard
 CC recombinant methodology. Zace2 polypeptides are useful for treating an
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
 CC diseases associated with inflammation like arthritis and enterocolitis,
 CC as targets for identifying modulators of zinc protease activity, for
 CC screening or identifying new angiotensin-converting enzyme (ACE)
 CC inhibitors, and as a basis for rational drug design for inhibitory
 CC molecules. The nucleic acids can be used to detect the expression of a
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and
 CC for detecting and localizing Zace2 gene expression in tissue samples,
 CC to determine whether a subject's chromosomes contain a mutation in the
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.
 CC Inhibitors of ACE are used for treating hypertension of various
 CC conditions, including left ventricular systolic dysfunction, progressive
 CC renal impairment, scleroderma renal crisis, congestive heart failure due
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
 CC used to treat infertility while Zace2 antagonists are used for inducing
 CC infertility. The present sequence represents a cDNA encoding the human
 CC Zace2 protein.
 CC
 SQ Sequence 3334 BP; 1011 A; 640 C; 754 G; 929 T; 0 other;
 XX
 XX
 Alignment Scores:
 Pred. No.: 0 Length: 3334
 Score: 4291.00 Matches: 805
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-978-385-2 (1-805) x AAC84366 (1-3334)
 QY 1 MetSerSerSerSerTrpLeuLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
 Db 35 ATGTCAAGCTCTTCCCTGCTCTTCCACGCTTGTGCTGTAACGTGCTGATGCTCACAC 94
 QY 21 IleGluGluGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
 Db 95 ATTAGAGCAACAGGCCAAGACATTTTGGACAAAGTTTAAACAGAGAGCCGAGACCTGTTT 154
 QY 41 TYRGLSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsnValGln 60
 Db 155 TATCAAAAGTCTACTTCTTGTGAAATTTAAACACCAATATTACGAGAGAAATGTCCAA 214
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80

Db 215 AACATGAATATGCTGGGGCAAAATGCTGCTTTTAAAGAAACAGTCCACACTTGGC 274
QY 81 GlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
Db 275 CAATGTATCCACTACAGAAATTCAGAAATTCACAGTCAAGCTTCACTGAGGCTCTCT 334
QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
Db 335 CACCAAAATGGTCTTCAAGTCTCAGAGCAACAGCAAGCAAGCGTTGAACCAATTCGA 394
QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
Db 395 AAATACATGAGCACCACTACACTACTGCAAAAGTTTGAACCCAGATATATCCACAAGA 454
QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
Db 455 TGGTTTATCTTGAACCAAGGTTTGAATGAATATATGCAACAGCTTATGACTACATGAG 514
QY 161 ArgLeuTyrAlaThrPheLysSerThrArgSerGluValGlyLysGlnLeuArgProLeuTyr 180
Db 515 AGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGTGGGCAAGCAGCTGAGCCATTATAT 574
QY 181 GlnGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAspTyrGly 200
Db 575 GAAGAGTATGTGCTTGAATAATGAGATGCAAGAGCAAAATCATATGAGACTATGAGG 634
QY 201 AspTyrTyrArgGlyLysPyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
Db 635 GATATATGGAGAGACACTATGACATTAATGGGATGAGTGGCTATGACTACAGCCGGC 694
QY 221 GlnLeuIleGluAspValGlnHisThrPheGlnGluIleLysProLeuTyrGlnHisLeu 240
Db 695 CACTGATTTGAAGATGTGGAACATACCTTTGAAGAGATTAACCATATATGACATCTT 754
QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db 755 CATGCTATGTGAGGCAAGTGTGATGAAATGCTATTCCTATATCATGCTCAATTTGA 814
QY 261 CysLeuProAlaHisLeuLeuGlnLysAspMetTyrArgPheTyrThrAsnLeuTyrSer 280
Db 815 TGGCTCCCTGCATCTTGGTGTGATATGCTGGGAGTAACTTTTGGACAAATCTGACTCT 874
QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 875 TTGACAGTTCCTTGGACAGAAACAAACATAGATGATGATGCAATGTTGAGCCAG 934
QY 301 AlaTyrAspAlaGlnArgIlePheLysGlnAlaGluLysPhePheValSerValGlyLeu 320
Db 935 GCCTGGATGACAGAGATATTTCAAGAGGCCGAGAAATCTTGTATCTGTTGCTCT 994
QY 321 ProAsnMetThrGlnGlyPheThrPheGluAsnSerMetLeuThrAspProGlyAsnValGln 340
Db 995 CCTAATATGACTCAAGAGATTTCTGGGAAATTCATGCTACAGGACCCGGAATATTTAG 1054
QY 341 LysAlaValAlaCysHisProThrAlaThrAspLeuGlnLysGlyAspPheArgIleLeuMet 360
Db 1055 AAAGCAGCTCTGCCATCCACAGCTTGGGAGCTGGGGAAGGGGACTTCAGATCTTATG 1114
QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyHisIleGln 380
Db 1115 TGCACAAAGGTACATGAGACACTTCCGACAGCTCATCATGAGATGGGAGATTCAG 1174
QY 381 TyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsnGlyAlaAsnGlnLysPhe 400
Db 1175 TATGATATGCAATGCTGCACAACTTTTCTGCTAAGAAATGAGCTTAATGAGATTC 1234
QY 401 HisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
Db 1235 CATGAAGCTGTGGGAAATCATGTCACTTCTGCAGCCACCACTTAACCATTTAAATCC 1294
QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440
|||||

Db 1295 ATTGCTTCTGTGACCCGATTTTCAAGAAACATGAACAGAAATAACTTCCTGCTC 1354
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTyrArg 460
Db 1355 AAACAGACAGCTCAGCATTTGTTGGAGCTTGGCACTTACTTACATGATTAAGAAATGGAGG 1414
QY 461 TrpMetValPheLysGlyLysIleProLysAspGlnThrPheLysTyrPheGluMet 480
Db 1415 TGGATGGCTTTAAAGGGGAATTTCCCAAGACAGTGGATGAAAGAAAGTGGTGGAGATG 1474
QY 481 LysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCysAspPro 500
Db 1475 AAGCAGAGATAGTTGGGGTGGTGGAACTTCCCATGATGAAACATCTGAGACCCC 1534
QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520
Db 1535 GCATCTGTTCATAGTTTCTTAATGATTAATCTCATTTCAATGATTAACAGACCCCTT 1594
QY 521 TyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaLysHisGlnGlyProLeuHis 540
Db 1595 TACCATTTCCAGTTTCAAGAAAGCACTTTGTCAACGACTAAACATGAAGCCCTGTGCAC 1654
QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyLysLeuPheAsnMetLeuArgLeu 560
Db 1655 AAATGTGACATCTCAAACTTACAGAAAGCTGGACAGAAAGCTGTTCAATATGCTGAGGCTT 1714
QY 561 GlyLysSerGluProTyrThrLeuAlaLeuGlnLysAsnValGlyAlaLysAsnMetAsn 580
Db 1715 GGAATATCAGAACCTTGACCTTGACCTTGAATATTTGTAGAGGCAAAAGCAATGAAT 1774
QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPheLysAspGlnAsnLys 600
Db 1775 GTAAGGCCAGCTGCTCACTACTTATGAGCCCTTATTTACGTGGCTGAAGAACGAGAACAG 1834
QY 601 AsnSerPheValGlyTyrPheThrAspTyrPheProTyrAlaAspGlnSerIleLysVal 620
Db 1835 AATTCTTTTGGGATGAGTACCGACTGGAGTGCATATGACAGCCAAACATCAAAAGTG 1894
QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluThrPheAsnAspAsnGluMet 640
Db 1895 AGGATTAAGCTTAATATCACTCTTGGAGATTAACATATGAATGAAGCAATGAATGAAATG 1954
QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgIleTyrPheLeuLysValLysAsn 660
Db 1955 TACTGTTCCGATCATCTGTTGCTATGCTATGAGGAGTACTTTTAAAGTAAATTAAT 2014
QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680
Db 2015 CAGATGATTTCTTTTGGGAGAGAGATGTGCAAGTGGCTTAATTTGAACCAAGATCTCC 2074
QY 681 PheAsnPhePheValThrAlaProLysAsnValSerIleIleProArgThrGluVal 700
Db 2075 TTTATATTTCTTGTCACTGCACTTAATAAATGTGCTGATATCAATCTTCAAGCTGAATG 2134
QY 701 GlnLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2135 GAAAGGCCAATCAGAGATGCTCCGAGCGCTATCATATGCTTCCGCTGATGACAAAC 2194
QY 721 SerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2195 AGCTTCAAGTTTCTGGGGAATACAGCCAAACCTTGGAACCTTCAACAGCCCTGTTTCC 2254
QY 741 IleTyrPheLeuIleValPheGlyValValMetGlyValIleValValGlyIleValIleLeu 760
Db 2255 ATATGCTGATTTGTTTGGAGTTTGATGGAGTGTATGTTGTCATGTGCAATCTG 2314
QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780
Db 2315 ATCTTCACTGGAGTACAGATTCGGAAGAAATAAAGCAAGAGTGGAGAAATTCCT 2374
QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800
Db 2375 TATGCCCTCATCATATTTACCAAAAGGAAATAATATCCAGAGATTCAAAAAAGCTGATGAT 2434
|||||

QY 341 lysalalalcyshsptromralatrpaspseuglysglyaspheargileumet 360
 |||||
 Db 1102 AAAGACTGCGATCCACACAGCTTGGACCTGGGAAGGCGACTTCAGGATCCTTATG 1160

QY 361 CysThrLysValThrMetAspSerPheLeuThrAlaHisHISGLMetGlyHisIleGln 380
 |||||
 Db 1162 TGCACAAAGGTGACAAATGGACACTTCTGCAGAGCTCATCATGAGATGGGCATATTCAG 1221
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluPhe 400
 1222 TATGATATGGCATATGCTGCACAAACCTTTTCTGTAAGAAATGGAGTAATGAAGATTC 1281
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 1282 CATGAAGCTGTGGGAAATCATGTCACTTCTGCAGCCACACCTTAACATTAAATCC 1341
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440
 1342 ATTGCTTCTGTCACCCGATTTTCAGAGACAAATGAAACGAAATTAATCTCTGCTC 1401
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTyrArg 460
 1402 AAACAAGCAGCTCAGATTGTGGAGCTGCGCATTTACTTACTTACATGTAGAGAGTGAGG 1461
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTyrTrpGluMet 480
 1462 TGGATGCTCTTAAAGGGGAAATCCCAAGACAGTGGATGAAAGTGGTGGAGATG 1521
 QY 481 LysArgGluIleValGlyValGluProValProHisAspGluThrTyrCysAspPro 500
 1522 AAGCGAGAGATAGTTGGGGTGGTGGAGACTGTGCCCCATGATGAACATACGTGATCCCC 1581
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520
 1582 GCATCTCTGTCATGTTCTTAATGATTAATCATTCATTCGATTCGATTAACAAGACCCCTT 1641
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaIleLysHisGluGlyProLeuHis 540
 1642 TACCAATTCAGATTCCAAAGACACTTGTCAAGCAGCTAAACAGAAAGGCCCTGTGCAC 1701
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560
 1702 AATATGACATCTCAAACTCTACAGAGCTGCAGCAAGAACTTCAATATGTGTAGGCTT 1761
 QY 561 GlyLysSerGluProTyrPheLeuAlaLeuGluAsnValGlyAlaLysAsnMetAsn 580
 1762 GGAATATCAGAACCTGCAGCCTAGCATGTGAAATGTGTGAGAGCAAGACATGAAT 1821
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrLeuLysAspLysAsnLys 600
 1822 GTAAAGCCCACTGCTCAACTACTTGTGAGCCCTTAATTAATCTGCTGTAAGACCAAG 1881
 QY 601 AsnSerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIleLysVal 620
 1882 AATTTCTTTGTGGATGGAGTACCGACTGAGTCCATATGCAAGACCAAGATCAAAATG 1941
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGlyTrpAsnAspAsnGluMet 640
 1942 AGGATTAAGCCTAAATCACTCTTGTGAGATTAAGCATATGAATGAAGACCAATGAATG 2001
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660
 2002 TACCTGCTGCATCTGTGCAATGCTATGAGGCACTACTTTTAAAGTAAATAAT 2061
 QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680
 2062 CAGATGATCTTTTGGGAGAGGAGATGTGAGAGTGGCAATTTGAACCAAGATTCCTCC 2121
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700
 2122 TTAAATTTCTTGTCACTCCACTCAAAATGTGTGTGATATCATCTTAAGAACTCAAGTT 2181
 QY 701 GluLysAlaIleLeuMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
 2182 GAAAGGCAATCAGATGTCCGAGACCGATATCAATGATGTCTTCCGTGTGATGACAC 2241

QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
 |||||
 Db 2242 AGCCTAGATTTCTGGGGATACAGCAACACTTGGACCTCTTAACCAAGCCCTGTTTCC 2301
 QY 741 IleTrpLeuIleValPheGlyValAlaMetGlyValIleValValGlyIleValIleLeu 760
 2302 ATATGCTGATTTGTTTGGAGTTGTCATGAGTGAATGATGTTGTCATGTGATCTG 2361
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnPro 780
 2362 ATCTTCACTGGATCAGATATCGAGAGACAAATAATTAACCAAGAGTGAGAAATCTCT 2421
 Db 2422 TATGCTTCATCGATATATTAACAAGGAAATTAATCCAGATTCACAAACACTGATGAT 2481
 QY 801 ValGlnThrSerPhe 805
 |||||
 Db 2482 GTTCAGACCTCCTTT 2496
 RESULT 4
 AAD02758 standard; cDNA; 3396 BP.
 ID AAD02758; standard; cDNA; 3396 BP.
 AC AAD02758;
 DT 31-MAY-2001 (first entry)
 XX Human angiotensin converting enzyme-2 (ACE-2) cDNA.
 DE Human angiotensin converting enzyme-2 (ACE-2) cDNA.
 XX Human; angiotensin converting enzyme-2; ACE-2; peptidyl dipeptidase A;
 KW screening; therapy; hypertension; congestive heart failure; CHF;
 KM inflammation; pain; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 FH Key
 FT 82..2499
 FT CDS
 FT /tag= a
 FT /product= "Human angiotensin converting enzyme-2
 (ACE-2)"
 FT /EC_number= "3.4.15.1"
 FT /note= "This region is referred as SEQ.ID.NO.3 and is
 specifically claimed in claim 26"
 FT sig_peptide
 FT 82..135
 FT /*tag= b
 FT mat_peptide
 FT 136..2496
 FT /*tag= c
 FT /product= "Human mature angiotensin converting enzyme-2
 (ACE-2)"
 PN US6194556-B1.
 XX 27-FEB-2001.
 PD 11-DEC-1997; 97US-0989299.
 PF 11-DEC-1997; 97US-0989299.
 PR 11-DEC-1997; 97US-0989299.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA Acton St., Robison KE;
 PI WPI: 2001-210604/21.
 DR P-PSDB; AAY72667.
 XX Novel genes encoding angiotensin converting enzyme-2 useful as
 PT antisense or antigene agents for therapeutics, diagnostics and
 PT screening assays -
 XX Claim 1; Fig 1; 76pp; English.
 PS The present sequence is human angiotensin converting enzyme-2 (ACE-2)
 CC

CC CDNA. ACE is also referred as peptidyl dipeptidase A. Nucleic acid
CC sequence encoding ACE-2 is useful as antisense or antigenic agents for
CC sequence specific modulation of gene expression or in the analysis of
CC single base-pair mutations in the gene. Nucleic acid sequence encoding
CC ACE-2 is useful in therapeutics, diagnostics and in screening assays.
CC ACE-2 antagonist is used to treat hypertension or congestive heart
CC failure (CHF). ACE agonist is used to reduce the inflammation and pain
CC resulting from an insect sting or bite, which was accompanied by an
CC injection of bradykinin. Anti-ACE-2 antibodies are used to monitor ACE-
CC protein levels for determining the disease or condition associated with
CC an aberrant protein level.

Sequence 3396 BP; 1034 A; 659 C; 772 G; 931 T; 0 other;

Alignment Scores:	
Pred. No.:	0
Score:	4291.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	22
Length:	3396
Matches:	805
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-978-385-2 (1-805) x AAD02758 (1-3396)

QY	1	MeberSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr	20
Dp	82	ATGCAAGCTCTTCTCGAGCTCTTCTCAGCCTTCTCTGTAAGCTGCTCACTCCACC	144
QY	21	IleGlnGlnAlaAlaLysThrPheLeuAspLysPheAsnHisGlnAlaGlnAspLeuPhe	40
Dp	142	ATTGAGGAAAGCGCCAGACACTTTTGGACAAAGTTTAACCCGAGAGCCGAGACTTTC	201
QY	41	TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGlnGlnAsnValGln	60
Dp	202	TATCAAGTTCACCTGCTTCTTGAAATTATACCCCAATTAATCTAGCAAGAAAGTCCAA	261
QY	61	AsnIleAsnAsnAlaGlnAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAla	80
Dp	262	AACATGAAATTAATGCTGGGGCAAAATGGCTCTCTTTTAAAGAAAGCTCCACACTTGGC	321
QY	81	GlnMetTyrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu	100
Dp	322	CAAAATGATATCCACACAAAGAAATTCAGAAATCTCAGTCAAGCTTCAAGCTCCAGGCTCTT	381
QY	101	GlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThrIleLeu	120
Dp	382	CAGCAAAATGGGCTTTCAGTCTGTCTCAGAAAGACAAGACAAGCGTTGAACAACAATTCTA	441
QY	121	AsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAsnProGlnGln	140
Dp	442	AATCAAAATGACACCAATCTACAGTACGCGAAAAAGTTTGTAACCCAGATTAATCCACAAGAA	501
QY	141	CysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGln	160
Dp	502	TGCTATTACTTGACACAGGTTTGAAATGAATTAATGGCAAAAGCTTTAGACTCAACAAGAG	561
QY	161	ArgLeuThrPalatrpGlnSerTrpArgSerGlnValGlyLysGlnLeuArgProLeuTyr	180
Dp	562	AGGCTCTGGGCTGGGAAGCTGAGATCTGAGGTGGGCAAGCAGCTGAGGCCATTATAT	621
QY	181	GlnGlnTyrValValIleLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAspTyrGly	200
Dp	622	GAAACAGATATGGCTTGAAAAAAGCAATGGCAAGACCAATCTATTATGAGGACTATGGG	681
QY	201	AspTyrTrpArgGlnAspTyrGlnValAsnGlyValAspGlyTyrAspTyrSerArgGly	220
Dp	682	GATTATTGGAGAGAGACTATGAAGTAATGGGGTAAATGAGCTATGACTACAGCCGGGCG	741
QY	221	GlnLeuIleGlnAspValGlnHisThrPheGlnGlnIleLysProLeuTyrGlnHisLeu	240
Dp	742	CAGTGTGATGAAGATGGGAACAATACCTTTGAAGAGAGTTAAACCATTAATATGAACAATCTT	801
QY	241	HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly	260

Dd	802	CATGCTATGTGAGGCGCAAAAGTTGATCAAGCCCTATCTCTCTATATCATGTCCAATTGGA	861
Qy	261	CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheThrAsnLeuTyrSer	280
Dd	862	TGCGTCCCTGCTCATTTTGCTTGGTGATATGTGGGGTAGATTTTGGACAAATCTGTACTCT	921
Qy	281	LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln	300
Dd	922	TTTGACAGTTCCCTTTGGACAGAAACCMAACATGATGATTACTGATCATGTCATGTGGACAG	981
Qy	301	AlaTrpAspAlaGlnArgGlyIlePheLysGluAlaGlnLysPhePheValSerValGlyLeu	320
Dd	982	GCTTGGGATGACAGAAATATTCAGAGGAGCCGGAAGAAGTCTTGTATCTTGTGGCTT	1041
Qy	321	ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln	340
Dd	1042	CCTATATGACTCAAGGATTCCTGGGAAATAATTCATGCTTAACGAGACCCAGAAATGTTCCAG	1101
Qy	341	LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet	360
Dd	1102	AAAGACAGTCTGCCATCCCAACAGCTTGGGACCTGGGGAAGGGCGACTTCAGATCCTATG	1161
Qy	361	CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetClyHisIleGln	380
Dd	1162	TGCACAAAGGTGACAAATGAGACGATCTTCAGCGTCATCATGAGATGGGGCATATCCAG	1221
Qy	381	TyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsnGlyValAsnGluIlePhe	400
Dd	1222	TATGATATGGCATATGCTGTGCACAACTTTTCTCTTAAGAAATGGAGCTATATAGAGATTC	1281
Qy	401	HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer	420
Dd	1282	CATGACGCTGTTGGGGAATCATGTCACTTTCGACCCACACCTACAGCATTTAAATCC	1341
Qy	421	IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu	440
Dd	1342	ATTGGTCTCTGTCACCCGATTTTCAAGAGACAAATGAAGAAACGAAATAACTTCTGTCTC	1401
Qy	441	LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg	460
Dd	1402	AAACACAGCACCTCAGATGTTGGGACTCTGCCATTTACTTACTGATGTAGAGAGTGGAGG	1461
Qy	461	TrpMetValPheLysGlyGluIleProLysAspGluTrpMetLysLysTrpTrpLeuMet	480
Dd	1462	TGATGTGCTTTAAAGGGGAATTCCTCAAAAGACAGTAGATGAAAGAGTGGGAGATG	1521
Qy	481	LysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCysAspPro	500
Dd	1522	AAGGAGAGATAGTTGGGGTGGTGGAACTGTGCCCATGATGAAGAACATACGTATCCCC	1581
Qy	501	AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu	520
Dd	1582	GCACTCTGTTCATGTTCTTAATGATATCTCATTTATTCGATTAATTCACAAGGACCTTT	1641
Qy	521	TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGlyGlyProLeuHis	540
Dd	1642	TACCAATTCAGTTCAGAAGACACTTTGTCAACGACGCTAAACATGAAAGGCCCTGTGCAC	1701
Qy	541	LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu	560
Dd	1702	AAATGTACATCTTCAAACTCTACAGAGCTGGCAGCAAAACTGTCATATGCTGTAGCTT	1761
Qy	561	GlyLysSerGluProTrpThrIleAlaIleGluAsnValValGlyValLysAsnMetAsn	580
Dd	1762	GGAAATATCAGAACCCCTGGACCCCTAGGACTTGGAAAATGTTGTGGAGCAAGAACATGAT	1821
Qy	581	ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys	600
Dd	1822	GTAAGCCACTGCTCAACTACTTTGAGCCCTTATTTACTGCTGTAAGACACAGACACAG	1881
Qy	601	AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal	620

QY	401	HisGlnAlaValGlyGluIleMetSerIleSerAlaAlaThrProLysHisIleLysSer	420
Db	1282	CATTAAGCTGTGGGGAAATCAATGACTACTTTCGACGCCACCTTAAGCATTTAAATACC	1341
QY	421	IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu	440
Db	1342	ATTGGTCTCTGTACCCGATTTTCAAGAAAGCAATGAAACAGAAATAAACTCTCTGCTC	1401
QY	441	LysGlnAlaLeuThrIleValIleGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg	460
Db	1402	AAACAACACACTCAAGATTTGTGGGACTGCGCCATTACTTACATGTTAGAGAAAGTGAGG	1461
QY	461	TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpGlnMet	480
Db	1462	TGGATGGTCTTTAAAGGGGAAATTCCTCAAGACCACTGATATAAAGAGGGGAGATG	1521
QY	481	LysArgGluIleValIleGlyValIleGluProValProHisAspGluThrTyrCysAspPro	500
Db	1522	AAGCAGAGATAGTGGGGGTGGTGGAACTGTGCCCCATGATGAACATACTGTGACCCC	1581
QY	501	AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu	520
Db	1582	GCATCTCGTGTCCATGTTTCTTAATGATATCATTCATTCGATATTAACAAGACACCTT	1641
QY	521	TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaIleLysHisGlnGlyProLeuHis	540
Db	1642	TACCAATTCACGTTTCCAAAGAGCACTTGTCAAGCAGCTAAACATGAAAGCCCTCGAC	1701
QY	541	LysCysAspIleSerAsnSerThrGluIleGlnIleLysLeuPheAsnMetLeuArgLeu	560
Db	1702	AAATGTGACATCTCAAACTGTACAGAGCTGGACACAAACTGTCAATATGCTGAGGCTT	1761
QY	561	GlyLysSerGluProTrpThrIleAlaLeuGluAspValIleGlyAlaLysAsnMetAsn	580
Db	1762	GGAAATATCAGACACCTGGACCTAGCATTTGGAAATGTTGTAGAGACAAGACATGAT	1821
QY	581	ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys	600
Db	1822	GTAAGCGCACTGCTCAACTCTTTGAGCCCTTATTACTGCGCTAAAGCACCAACAG	1881
QY	601	AsnSerPheValIglyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal	620
Db	1882	AAATCTTTTGGGATGAGATACCGCATGGAGTCCATATGACACCAACCACTCAACTG	1941
QY	621	ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGlnMet	640
Db	1942	AGGATAGACCTTAAATATAGCTCTTGGAGATAAAGCATATGATGACACACATGAATG	2001
QY	641	TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn	660
Db	2002	TACCTGTTCGATCATCTGTGTCATATGCTATGAGCGCATCTACTTTTAAATATAAAAT	2061
QY	661	GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer	680
Db	2062	CAGATGATTTCTTTTGGGGAGGAGATGTGCGGATGGCTAAATTTGAACCAAGATCTCC	2121
QY	681	PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal	700
Db	2122	TTTATTTCTTTGTGCACTGCACCTAAATATGTCTGATATCATTTCTTGAAGTGAATT	2181
QY	701	GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn	720
Db	2182	GAAGAAGGCATACAGATGTCCCGGAGCGATCAATGATGCTTCCGTCGTGATGATCAAC	2241
QY	721	SerLeuGluPheLeuGlyIleGlnProThrIleGlyProProAsnGlnProProValSer	740
Db	2242	AGCCTAAGATTTCGGGGATACAGCAACACTTGGACCTCTTACACAGCCCCCTGTGTC	2301
QY	741	IleTrpLeuIleValPheGlyValAlaMetGlyValIleValIleValGlyIleValIleLeu	760
Db	2302	ATAAGCTGATGTGTTTGGAGTGTGATGAGGAGTATATGTCGTGGCATTTGTATCTCTG	2361
QY	761	IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnPro	780

Db 622 GAAGATATGTGCTTGGAAAAATGAGATGCGAAGACCAATCATATGAGACTATGG 681
 QY 201 ASPTTTPATGAGLAsPTyrgLValAsnGLYValAspGLYTYrAsPTySArGLY 220
 Db 662 GATTATTTGGAGAGAGACTATGAAATGAGTGGGTAGTGGCTATGACTACAGCGGCGC 741
 QY 221 GlnLeuIleGLuAsPValGLuHISThrPheGLuGLuIleLysProLeuTYrGLuHISLeu 240
 Db 742 CAGTGATGGAAGATGAGATGCAATACCTTGAAGAGATTAACCATATATGAAACATCTT 801
 QY 241 HisAlaTYrValArgAlaLysLeuMetAsnAlaTYrProSerTYrIleSerProIleGLY 260
 Db 802 CATGCCATGTGAGGGCAAAAGTTGATGATGCTATCCCTCCATATCAGTCAATGGA 861
 QY 261 CysLeuProAlaHISLeuGLYAspMetProLysArgPheTYrPheTYrAsnLeuTYrSer 280
 Db 862 TGCCCTCCCTGCTCATTTGCTTGTGATATGTTGGGGTAGATTTTGGCAAAATCTGTCTCT 921
 QY 281 LeuThrValProPheGLYGLuLysProAsnIleAspValThrAspAlaMetValAspGLn 300
 Db 922 TTGACAGTTCCTTGGAGCAAAACCAACATATGATGTACTGATGCAATGGTGGACGAG 981
 QY 301 AlaTrpAspAlaGLuArgIlePheLysGLuAlaGLuLysPhePheValSerValGLYLeu 320
 Db 982 GCCTGGGATGCAAGAGATATTCAGAGAGGCGAGAAAGTTCTTGTATCTGTGGTCTT 1041
 QY 321 ProAsnMetThrGLnGLYpHeTrpGLuAsnSerMetLeuThrAspProGLYAsnValGLn 340
 Db 1042 CCTATATGACTCAAGGATTTCTGGGAAATTCATGCTAACGCGCCAGGAAAGTTTCAG 1101
 QY 341 LysAlaValCysHISProThrAlaTrpAspLeuGLYLysGLYAspPheArgIleLeuMet 360
 Db 1102 AAACAGACTGCCATCCACAGCTGGGACCTGGGGAAGGCGCACTTCAGATCTTATG 1161
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHISISGLuMetGLYHISLeGLn 380
 Db 1162 TGCAACAAGGTGACAAATGAGAGACTTCTGCACAGCTCATATGAGATGGGCAATATCCAG 1221
 QY 381 TyrAspMetAlaTYrAlaAlaGLnProPheLeuLeuArgAsnGLYAlaAsnGLYpHe 400
 Db 1222 TATATATGCGATATGCTGCGCAACCTTTTCTGTCTAAGAAATGAGGTAATGAAAGATTC 1281
 QY 401 HisGLuAlaValGLYLeuIleMetSerLeuSerAlaAlaThrProLysHISLeuLysSer 420
 Db 1282 CATGAACCTGTTGGGAAATCATGTCTTCTGACAGCCACACCTAAGCATTAATAATCC 1341
 QY 421 IleGLYLeuLeuSerProAspPheGLnGLuAspAsnGLuThrGLuIleAsnPheLeu 440
 Db 1342 ATTGTCTTGTCTACCCGATTTTCAAGAAACAAATGAAACAGAAATTAACCTTCTGCTC 1401
 QY 441 LysGLnAlaLeuThrIleValGLYThrLeuProPheThrTYrMetLeuGLYLeuSTPArg 460
 Db 1402 AAACAACACACTCAGATTTGGGACCTCTGCCATTTACTTACATGTTAGGAAGTGGAGG 1461
 QY 461 TrpMetValPheLysGLYLeuIleProLysAspGLnTrpMetLysLysSTPTrpGLuMet 480
 Db 1462 TGGATGCTCTTAAAGGGGAAATCCCAAGACAGAGATGAAAGAAAGTGGTGGAGATG 1521
 QY 481 LysArgGLuIleValGLYValAlaGLuProValProHisAspGLuThrTYrCysAspPro 500
 Db 1522 AAGGGAAGAGATGTTGGGGTGGGAGACCTGCTCCCATATGATTAACATCTCTGTGACCCC 1581
 QY 501 AlaSerLeuPheHISValSerAsnAspTYrSerPheIleArgTYrTYrThrArgThrLeu 520
 Db 1582 GCATCTCTGTTCATGTTTCTAATGATTAATCAATCAATTTGATATTAACAGAGACCTT 1641
 QY 521 TyrGLnPheGLnPheGLnGLuAlaLeuCysGLnAlaAlaLysHISGLuGLYProLeuHIS 540
 Db 1642 TACCAATTCAGATTCAAGAGACCTTGTCAAGACAGCTTAACATGTAAGGCCCTGTGAC 1701
 QY 541 LysCysAspIleSerAsnSerThrGLuAlaGLYGLnLysLeuPheAsnMetLeuArgLeu 560
 Db 1702 AATATGACATCTCAAACTCTACAGAGAGCTGAGCAAGAAACTGTTCATATATGCTGAGGCTT 1761

QY 561 GLYLysSerGLuProTrpPheLeuAlaLeuGLuAsnValValGLYAlaLysAsnMetAsn 580
 Db 1762 GGAATAATCAGAACCCCTGAGACCTTGAATGGAATAATGTTGTAGAGCAAAAGCAATGAT 1821
 QY 581 ValArgProLeuLeuAsnTYrPheGLuProLeuPheThrTrpLeuLysAspGLnLys 600
 Db 1822 GTAAAGCCACTGCTCAACTACTTGTGAGCCCTTATTTACTGCTGCTGAAGACCAAGCAAG 1881
 QY 601 AsnSerPheValGLYTrpSerThrAspTrpSerProTYrAlaAspGLnSerIleLysVal 620
 Db 1882 AATTCCTTTGTGGATGAGGTACCGACTGGAGTCCATATGACAGCAACCAAGATCAAGTG 1941
 QY 621 ArgIleSerLeuLysSerAlaLeuGLYAspLysAlaTYrGLuTrpAsnAspAsnGLuMet 640
 Db 1942 AGGATTAAGCTTAATAACAGCTCTTGGAGATTAAGCATATGAAACCAATGCAATGAAATG 2001
 QY 641 TyrLeuPheArgSerSerValAlaTYrAlaMetArgGLnTYrPheLeuLysValLysAsn 660
 Db 2002 TACCTGTTCCGATCATCTGTTCATATGCTATGAGGAGAGCTACTTTTAAAGTTAAATAAT 2061
 QY 661 GlnMetIleLeuPheGLYGLuGLuAspValArgValAlaAsnLeuLysProArgIleSer 680
 Db 2062 CAGATGATTCCTTTTGGGAGAGATGTCGAGTGGCTAATTTGAAACCAAGATCTCC 2121
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGLuVal 700
 Db 2122 TTTAAATTTCTTTGTCACTCCACTTAATAAATGTTGTGATATCATCTCTGAACCTGAAGTT 2181
 QY 701 GLuLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
 Db 2182 GAAAGGCCATCAGAGATGTCGCCGAGCCGTATCATATGATGCTTCCGTGTGATGACAAAC 2241
 QY 721 SerLeuGLuPheLeuGLYIleGLnProThrLeuGLYProProAsnGLnProProValSer 740
 Db 2242 AGCCTAGATTTCTGGGATACAGCCACCACTTGGACCTCTTAACAGGCCCTCTTTCC 2301
 QY 741 IleTrpLeuIleValPheGLYValAlaMetGLYValIleValValGLYIleValIleLeu 760
 Db 2302 ATATGGCTGATGTTTGTGAGATTTGATGGAGAGATAGTGGTGGCATTTGTCATCTCG 2361
 QY 761 IlePheThrGLYIleArgAspArgLysLysLysAsnLysAlaArgSerGLYGLuAsnPro 780
 Db 2362 ATCTTACAGCGCATCGATCGAATCGGAAGAAATAAATTAAGCAAGAGTGGAGAAAATCT 2421
 QY 781 TyrAlaSerIleAspIleSerLysGLYGLuAsnAsnProGLYpHeGLnAsnThrAspAsp 800
 Db 2422 TATGCTCCATCATGATTTAGCAAAAGAGAAATAATCCAGATTCCAAAAACCTGATGAT 2481
 QY 801 ValGLnThrSerPhe 805
 Db 2482 GTTCAGACCTCCTT 2496

RESULT 7
 AAS21279
 ID AAS21279 standard; cDNA; 3732 BP.
 XX
 AAS21279;
 XX
 24-OCT-2001 (first entry)
 DE Human cDNA sequence encoding for PRO1885 polypeptide.
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
 KW breast; prostate; cervical; tumour necrosis factor- α lpha; TNF- α lpha;
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.
 OS Homo sapiens.
 XX
 PN WO200140466-A2.
 PD 07-JUN-2001.

XX 01-DEC-2000; 2000MO-US32678.
 XX
 PR 01-DEC-1999; 99MO-US28301.
 PR 01-DEC-1999; 99MO-US28634.
 PR 02-DEC-1999; 99MO-US28551.
 PR 02-DEC-1999; 99MO-US28564.
 PR 02-DEC-1999; 99MO-US28565.
 PR 09-DEC-1999; 99US-0170262.
 PR 16-DEC-1999; 99MO-US30095.
 PR 20-DEC-1999; 99MO-US30911.
 PR 20-DEC-1999; 99MO-US30999.
 PR 30-DEC-1999; 99MO-US31243.
 PR 06-JAN-2000; 2000MO-US00277.
 PR 06-JAN-2000; 2000MO-US00376.
 PR 11-FEB-2000; 2000MO-US03565.
 PR 18-FEB-2000; 2000MO-US04341.
 PR 18-FEB-2000; 2000MO-US04342.
 PR 22-FEB-2000; 2000MO-US04414.
 PR 24-FEB-2000; 2000MO-US04914.
 PR 24-FEB-2000; 2000MO-US05004.
 PR 01-MAR-2000; 2000MO-US05601.
 PR 20-MAR-2000; 2000MO-US07377.
 PR 21-MAR-2000; 2000MO-US07532.
 PR 30-MAR-2000; 2000MO-US08439.
 PR 17-MAY-2000; 2000MO-US13705.
 PR 22-MAY-2000; 2000MO-US14042.
 PR 30-MAY-2000; 2000MO-US14941.
 PR 02-JUN-2000; 2000MO-US15264.
 PR 10-NOV-2000; 2000MO-US30873.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Bereshti M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WL, Zhang Z;
 XX
 DR P-PSDB; AAU12207.
 XX
 PT WPI; 2001-408281/43.
 PT
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical
 XX
 PS
 XX Claim 3; Fig 71; 813pp; English.

XX
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 XX

Sequence 3732 BP; 1137 A; 722 C; 821 G; 1052 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 3732
 Score: 4142.00 Matches: 802

Percent Similarity: 85.33% Conservative: 1
 Best Local Similarity: 85.23% Mismatches: 2
 Query Match: 96.53% Indels: 138
 DB: 22 Gaps: 1

US-09-978-385-2 (1-805) x AAS21279 (1-3732)

QY 1 MetSerSerSerSerTrpLeuLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
 DB ATGTCAAGCTTCTTCTGGCTCTCTCTCAGCTTCTTCTGTAATGCTGCTCAGTCCACC 99
 QY 21 IleGluGlnAlaAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
 DB ATTGAGCAACAGCCCAAGACATTTTGGACAGTTTAAACCAAGTTTAAACCAAGCAAGCTGTTTC 159
 QY 41 TyrGlnSerSerLeuAlaSerTrpAsnThrTrpSerHisThrAsnIleThrGluGlnValGln 60
 DB TATCAAAAGTCACTTCTTCTTGAATTAATTAACACCAATTAATTAATTAATTAATTAATTAATTA 219
 QY 61 AsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAla 80
 DB AACATGATTAATGCTGGGAGCAAAATGCTCTCTTTTAAAGAACAGTCCACACTTGGC 279
 QY 81 GlnMetTrpProLeuGlnIleGlnAsnLeuThrValLysLeuGlnLeuAlaLeu 100
 DB CAAATGATCCACTACAAAGAAATTCAGAACTCCAGCAAGCTTCAAGCTGCAAGCTCTT 339
 QY 101 GlnGlnAsnGlySerSerValLeuSerGlnLysPheLysSerLysAspLeuAsnThrIleLeu 120
 DB CAGCAAAATGGCTCTCTGCTCTCTCAGAAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 399
 QY 121 AsnThrMetSerThrIleTrpSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
 DB AATTAATGAGACACCATCTACAGTACAGTAAAGTTTGTAAACCAATTAATTAATTAATTAATTA 459
 QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTrpAsnGlu 160
 DB TGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 519
 QY 161 ArgLeuTrpAlaTrpGlnSerTrpArgSerGlnValGlyLysGlnLeuArgProLeuTrp 180
 DB AGGCTCTGGCTGGGAAAGCTGGAGATCTAGGCTGGCAAGCAAGCTGAGGCCATTATAT 579
 QY 181 GluGluTrpValValLeuLeuLysAsnGluMetAlaArgAlaAsnHisTrpGluAspTrpGly 200
 DB GAAGGATATGCTGCTTAAATAATGAGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 639
 QY 201 AspTrpTrpArgGlyAspTrpGluValAsnGlyValAspGlyTrpAspTrpSerArgGly 220
 DB GATTAATGAGAGAGAGACTATGAAGTAATGGGTAGTGTGCTATGACTACAGCCGCGGC 699
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTrpGluHisLeu 240
 DB CAGTTGATTAAGATGTGGAACATACCTTTAAAGATTAACCATTAATTAATTAATTAATTAATTA 759
 QY 241 HisAlaTrpValArgAlaLysLeuMetAsnAlaTrpProSerTrpIleSerProIleGly 260
 DB CATGCCCTATGAGAGGCAAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 819
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpTrpAsnLeuTrpSer 280
 DB TGCCCTCCCTGCTCAATTCCTTGGTCAATVAGGGGTGATTTTGGCAAAATCTGACTCT 879
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
 DB TTGACAGTCTCTTTGGGACAGAAACCAACATAGATGTATGATGATGATGATGATGATGATGAT 939
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320
 DB GCCTGGATGACAGACAGATATTCAGAGAGCCGCAAGATGCTTGTATGCTGTTGCTT 999
 QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAspValGln 340

Db 1000 CCTAATATGACTCAAGATTCTGGAAAAATTCATGCTAACGAGACCAGAAATGTTTCAG 1059
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleuMet 360
 Db 1060 AAGAGCTGTGCATCCACAGCTGGGAGCTGGGAGAGGGGAGCTTCAGGATCTCTTAG 1119
 QY 361 CysThrLysValIleuMetAspPheLeuThrAlaHisIleGlyMetIleHisIleGln 380
 Db 1120 TGCACAAAGGTGACATGACAGACCTCCGACAGCTCATCAGATGGGGCATATCCAG 1179
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyPhe 400
 Db 1180 TATGATATGTCATATGCTCCACACACTTTCCTAAGAAATGAGCTAATAGAGATTC 1239
 QY 401 HisGlnAlaValGlyIleuMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 Db 1240 CATGAAAGCTGTGGGAAATCATGTCACTTCTGCAAGCCACCTAAGCATTTAAATCC 1299
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440
 Db 1300 ATTGGTCTCTGTACCCGATTTTCAGAGACATGAAACAGAAATAACTCTCTGCTC 1359
 QY 441 LysGlnAlaLeuThrIleValGlyIleuProPheThrTyrMetLeuGluLysTrpArg 460
 Db 1360 AACAAGACACTCAGATGTGGGACTGTGCATTTACTTACATGTTAGAGAGTGAGAG 1419
 QY 461 TrpMetValPheLysGlyIleuLeuProLysAspGlnTrpMetLysLysTrpIleuMet 480
 Db 1420 TGGATGGCTTTTAAAGGGGAAATTCACCAAGCCAGTGGATGAAAATGGTGGAGATG 1479
 QY 481 LysArgGluIleValGlyValGlyValGluProValProHisAspGluThrTyrCysAspPro 500
 Db 1480 AAGCAGAGATATGTTGGGGTGGTGAACCTGTGCCCATGTGTAACATGACTGTGACCCC 1539
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520
 Db 1540 GCATCTCTGTCCATGTTTGCATGATATACATTCATTCATCATATTAACAGAGCCCTT 1599
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaIleValHisGlyIleProLeuHis 540
 Db 1600 TACCAATTCAGATTTCAGAAAGCACTTTGTCAACAGCAGCTTAAACATGAAAGGCCCTGTGAC 1659
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyIleGlnLysLeuPhe----- 555
 Db 1660 AATGTGACATCTCAAACTCTACAGAACTGGACAGAAACTGTT-GTAAAGAAATACCTCA 1718
 QY 555 ----- 555
 Db 1719 AATGTGAAACCTCTCTAGTATTCAGATTAATTAATTCATTTCCATGCTAGTTTATTTG 1778
 QY 555 ----- 555
 Db 1779 ATTCTTGTCTTAAAAAAGAAATTTATGGCTCAAAATGCTCCTCATTTACAAACCAA 1838
 QY 555 ----- 555
 Db 1839 CATTTAATTTGTGTGACAGAGAACTAGACCATACACAATTTGGGTGGCCACCTCTT 1898
 QY 555 ----- 555
 Db 1899 TTCTCCATCATTAATACAGCCCTCTTCTCTGTGTAATGGAAGAAAGGGGTTTAG 1958
 QY 555 ----- 555
 Db 1959 GGTGGAATATATCTGTAAATATGCAATTTCTTTCTTATCTCCAGAAACAAATTTAGCCAA 2018
 QY 555 ----- 555
 Db 2019 GTCAAAAGAGAAACCATATGATCATATATATATATATATATATATATATATATATAT 2078
 QY 556 -----AsnMetLeuArgLeuGlyLysSerGluPro 565
 Db 2079 AAGGCCCTGAACCCCTTTTTTTTGTGTAGCAATATGCTGAGGCTGGAATAATCAGAAACC 2138

QY 566 TrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeu 585
 Db 2139 TGGACCCCTAGCATTTGGAATAATGTTGTAGGAGCC-AAGAAACATGATGTAAAGCCCTCTC 2197
 QY 586 AsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGly 605
 Db 2198 AACTACCTTTGAGCCCTTATTTACTCGCTGAAAGACCAAGAAACAGAAATCTTTTGTGGGA 2257
 QY 606 TrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSerLeu 625
 Db 2258 TGGAGTACCGAGCTGGAGTCCATATGACAGACCAAGCATCAAGTGAAGATTAAGCCTTAA 2317
 QY 625 sSerAlaLeuGlyLysPylsAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPheArgse 645
 Db 2318 ATCAGCTCTTGGAGATTAAGCATATGAAATGAAAGCAACAAATGATGATCTGTTCCGATC 2377
 QY 645 rSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetIleLeuPhe 665
 Db 2378 ATCTGTTCATATGCTATGAGGCGATCTTTTAAAAATCAATCAGATGATCTTTT 2437
 QY 665 eGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePhe 685
 Db 2438 TGGGAGAGAGATGTGGAGTGGCTATTTGAAACCAAGATCTCTTTATTTCTTTGT 2497
 QY 685 lThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAlaIleArg 705
 Db 2498 CACTGCACCTAAATAATGTCTGTATATCATCTTACATGAGATGTAAGAGGATCAG 2557
 QY 705 gMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPhe 725
 Db 2558 GATGTCCCGAGCGATATCATATGATGCTTCCGTGTAATGACACAGCCCTAGAGTTCT 2617
 QY 725 uGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerIleTrpLeuIleVal 745
 Db 2618 GGGGATACACCAACACTGTGACCTCTTAACACGCCCTGTTCCATATGAGCTGATGCT 2677
 QY 745 lPheGlyValAlaMetGlyValIleValValGlyIleValIleLeuIlePheThrGly 765
 Db 2678 TTTTGGAGTTGTGATGGAGTGTATGTTGGTGCATTTGATCATCTTATCTTACCTGAGAT 2737
 QY 765 eArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnProTyrAlaSerIleAs 785
 Db 2738 CAGAGATCGGAAAGAAATAATTAAGCAAGATGAGAAATCCTTATGCGCTCATGCA 2797
 QY 785 pIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPh 805
 Db 2798 TATTAGCAAGAGAAATAATATCCAGATTCAAAACACTGATGATTCAGACCTCCTT 2857
 QY 805 e 805
 Db 2858 T 2858
 RESULT 8
 AAS14880
 ID AAS14880 standard; cDNA: 2920 BP.
 AC AAS14880;
 XX
 XX
 DT 20-DEC-2001 (first entry)
 XX
 XX
 DE Human cDNA encoding novel human protein NHP #1.
 XX
 KW Human; novel human protein; NHP; ss; antidiabetic; antirheumatic;
 KW antidiabetic; cytosolic; antidiabetic; antidiabetic; antidiabetic;
 KW neuroprotective; neurotrophic; antiparkinsonian;
 KW anti-human immunodeficiency virus; antistimatic; vasotrophic; cardiant;
 KW hypotensive; anorectic; antinfertility; neuroleptic; anticonvulsant;
 KW antinflammatory; immunosuppressive; cerebroprotective; antimicrobial;
 KW antiinflammatory; antiseborrheic; dermatological; vasoconstriction;
 KW immunomodulator; antidiabetic; dermatological; vasoconstriction;
 KW gastrointestinal disorder; cardiovascular disorder; hypertension;
 KW coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;


```

OY 302 PASPALAGLInrgIlePheLysGluAlaGluLysPhePheValSerValGlyLeuProAs 322
Db 935 GGATGACAGAGAAATATTCAAGAGAGCCGGAAGACTTCTTGATCTGCTGGCTTCTTAA 994
OY 332 nMethrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGlnLysAl 342
Db 995 TATAGACTCAAGGATTCGGGAAATTCATGCTAACGGACCCAGGAATGTTGAGAAAC 1054
OY 342 aValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMetCysTh 362
Db 1055 AGTGTCCATGCCACACCTTGGGACCTGGGGAAGGCGACTTCAGATCTTATGTGCAC 1114
OY 362 rLysValThrMetAspPheLeuThrAlaHisGluMetGlyHisIleGlnIleLys 382
Db 1115 AAAGGTACATGAGACACTTCTGACAGCTCATGATGATGGGACATATCCAGTATGA 1174
OY 382 pMetAlaTrpAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPheHisG 402
Db 1175 TATGGCATATGCTGCACAACTTCTGCTTAAGAAATGAGCCTAATGAGATTCATGA 1234
OY 402 uAlaValGlyLulIeMetSerLeuSerAlaIleThrProLysHisIleLysSerIleG 422
Db 1235 AGCGTGGGGAATCATGCTCATCTTCGACGACACCTTAAGCATTTAAATCCATTGG 1294
OY 422 yLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLysG 442
Db 1295 TCTTGTGTCCACCGATTTTCAAGAAACATGAAACAGAAATTAATCTCTGCTAAACA 1354
OY 442 nAlaLeuThrIleValGlyThrLeuProPheThrYrMetLeuGlyLysTrpArgTrp 462
Db 1355 AGACATCAGATGTTGGGACTCTGCTCATTTACTTACATGATTAAGAAAGTGGAGTG 1414
OY 462 tValPheLysGlyLulIeProLysAspGlnTrpMetLysTrpTrpGluMetLysAr 482
Db 1415 GGCCTTTAAAGGGGAAATCCCAAGACAGATGATGAAAAAGTGTGGAGATGAACG 1474
OY 482 gGluIleValGlyValGluProValProHisAspGluThrYrCysAspProAlaSe 502
Db 1475 AAGGATAGTGGGGTGGTGGGAACTGTCCTCCATGATATAACATCTGTGACCCGATC 1534
OY 502 rLeuPheHisValSerAsnAspTrpSerPheIleArgTrpYrThrArgTrpLeuArg 522
Db 1535 TCTGTTCACATGTTCTAATGATTAATCTCATCTTCATGATATACCAAGGACCTTTTCA 1594
OY 522 nPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHisLys 542
Db 1595 ATTCCAGTTTCAAGAGACATTTGTCAAGACAGCTAAACATGAAGCCCTGCAAAATG 1654
OY 542 sAspIleSerAsnSerTrpGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGly 562
Db 1655 TCACATTCCAAACTCTACAGAACTGGACAGAAACTGTTCAATVTGCTGAGGTTGGAAA 1714
OY 562 sSerLeuProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnVal 582
Db 1715 ATCAGAACCTGGAGCCCTAGCATTTGGAAGAAAGTTGTGAGGACAAAGACATGATGA 1774
OY 582 gProLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsn 602
Db 1775 GCCACTGCTCAACTTCTTGAGCCCTTATTAATCTGCTGAAAGACAGAAAGAAATTC 1834
OY 602 rPheValGlyTrpSerThrAspTrpSerProTrpAlaAspGlnSerIleLysValArg 622
Db 1835 TTTTGTGGATGGATGCCAGCTGAGATCCATATGACAGACCAAGCATCAAGTGGAGAT 1894
OY 622 eSerLeuLysSerAlaLeuGlyAspLysAlaTrpGluTrpAsnAspAsnGluMetYr 642
Db 1895 AAGCTTAATAACCTCTTGAGATTAAGCATATGATGAGACAGACATGAATGTACCT 1954
OY 642 uPheArgSerSerValAlaTrpAlaMetArgGlnTrpPheLeuLysValLysAsnGlu 662
Db 1955 GTTCGATCATCTTCTCATATGCTATGAGGACGACTCTTTTAAAGTAAATAATCATGAT 2014
OY 662 tIleLeuPheGlyGlnGluAspValArgValAlaAsnLeuLysProArgIleSerPhe 682

```

```

Db 2015 GATTCCTTTTGGGAGAGAGATGTCGAGCGCTAATTTGAACCAAGAAATCTCTTAA 2074
OY 682 nPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGly 702
Db 2075 TTTCTTGTCACTGACCTCAATAAATGCTGTGATATCATCTCAGAACTGAGTTGAAA 2134
OY 702 sAlaIleArgMetSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerIle 722
Db 2135 GGCCATCAGATGATCCCGGAGCCGTATCATATGATGCTTCCGTGAATGAGACAGCT 2194
OY 722 uGluPheLeuGlyTrpIleGlnProThrLeuGlyProProAsnGlnProProValSer 742
Db 2195 AGAGTTTCTGGGATACAGCCACACTTGGACCTCTTAACGACCCCGTTTCCATATG 2254
OY 742 pLeuIleValPheGlyValValMetLysValIleValValGlyIleValIleLeuIle 762
Db 2255 GCTGATTTGTTTGGAGTTGTGATGGAGTGAATGATGGCATTCATCTGATCTT 2314
OY 762 eThrGlyIleArgAspArgLysLys 770
Db 2315 CACTGGGATCAGAGATCGAAGAG 2339

RESULT 9
ID AAS14890 standard; cDNA: 2911 BP.
AC AAS14890;
XX
XX
XX 20-DEC-2001 (first entry)
DE Human cDNA encoding novel human protein NHP #11.
XX
XX Human; novel human protein; NHP; ss; antidiabetic; antirheumatic;
XX antiarthritic; cytosolic; antilisteriosclerotic; vulnerrary;
XX neuroprotective; nootropic; antiparkinsonian;
XX anti-human immunodeficiency virus; antilastimatic; vasotropic; cardiant;
XX hypotensive; anorectic; antinfertility; neuroleptic; anticonvulsant;
XX antinflammatory; antibacterial; cerebroprotective; antimicrobial;
XX immunomodulatory; antidiabetic; dermatological; vasoconstriction;
XX gastrointestinal disorder; cardiovascular disorder; hypertension;
XX coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;
XX cachexia; male infertility; impotence; testicular cancer; lung tumour;
XX hyperproliferative disorder; pulmonary system disorder;
XX central nervous system disorder; bone disorder;
XX neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
XX Huntington's disease; schizophrenia; mania; dementia; paranoia;
XX panic disorder; learning disability; amyotrophic lateral sclerosis;
XX psychosis; autism; sleep disorder; immune system disorder;
XX Hashimoto's thyroiditis; musculo-skeletal system disorders;
XX multiple sclerosis; ischaemic brain injury; stroke; infectious disease;
XX diabetes mellitus; immunological disorder; asthma; AIDS;
XX acquired immunodeficient syndrome; leukaemia; rheumatoid arthritis;
XX inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;
XX neural system disorder; respiratory disorder; olfactory disorder;
XX wound healing.
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 213..998
XX FT /*tag= a
XX FT /product= "NHP #11"
XX
XX PD WO200174896-A1.
XX
XX 11-OCT-2001.
XX
XX 02-APR-2001; 2001MO-US10542.
XX
XX 03-APR-2000; 2000US-194118P.
XX
XX 29-SEP-2000; 2000US-236384P.
XX

```


FH Key Location/Qualifiers
 FT 106..2523
 FT /*Tag= a
 FT /product= "Zace2-5"
 FT /note= "the coding fragment is specifically claimed for"
 XX
 PN W0200070032-A1.
 XX
 PD 23-NOV-2000.
 XX
 PE 03-MAY-2000; 2000WO-US11932.
 XX
 PR 13-MAY-1999; 99US-0311482.
 PR 27-AUG-1999; 99US-0384706.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Pladdington CS, Petrie CR, Shoemaker KE, Bishop PD;
 XX
 DR WPI; 2001-025018/03.
 DR P-PSDB; ABA48097.
 XX
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 PT associated with inflammation such as arthritis and enterocolitis -
 XX
 PS Claim 10; Page 104-109; 125pp; English.
 XX
 CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood
 CC pressure regulation and fertility. Zace2 can be expressed by standard
 CC recombinant methodology. Zace2 polypeptides are useful for treating an
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
 CC diseases associated with inflammation like arthritis and enterocolitis,
 CC as targets for identifying modulators of zinc protease activity, for
 CC screening or identifying new angiotensin-converting enzyme (ACE)
 CC inhibitors, and as a basis for rational drug design for inhibitory
 CC molecules. The nucleic acids can be used to detect the expression of a
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and
 CC for detecting and localizing Zace2 gene expression in tissue samples,
 CC to determine whether a subject's chromosomes contain a mutation in the
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.
 CC Inhibitors of ACE are used for treating hypertension of various
 CC conditions, including left ventricular systolic dysfunction, progressive
 CC renal impairment, scleroderma renal crisis, congestive heart failure due
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
 CC used to treat infertility while Zace2 antagonists are used for inducing
 CC infertility. The present sequence represents a cDNA encoding the mouse
 CC Zace2-5 protein.
 XX
 SQ Sequence 2638 BP; 802 A; 556 C; 611 G; 669 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 0 Length: 2638
 Score: 3579.00 Matches: 661
 Percent Similarity: 89.578 Conservative: 60
 Best Local Similarity: 82.118 Mismatches: 84
 Query Match: 83.418 Indels: 0
 DB: 22 Gaps: 0
 US-09-978-385-2 (1-805) x AAC64368 (1-2638)
 QY 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
 DB 106 ATGTCCAGGCTCCCTCGGCTCTTCAGCTTGCTGTACTGCTAGTCCGCTC 165
 QY 21 ILGLIUGLInAlaLysThrPheLeuAspLysPheAsnHisGlnGluAspLeuPhe 40
 DB 166 ACCGAGGAAATATCCCAAGACATTTTAAACAACTTAAATCAAGAACTGAGCCGCTG 225
 QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGlnGluAsnValGln 60
 DB 226 TATCAAGAGTTCACCTGCTTCTTGGAATTAATTAATCACTTACTGAGAAATGCCCCA 285

QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
 DB 286 AAGATGAGTAGGCTGCTACACCCCAAAATGCTGCTTTTAAAGAACAGTCAAGACTGCC 345
 QY 81 GlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuAlaLeu 100
 DB 346 CAAAGTTCTCAGCTACCAAGAAATCCGAGCTCCGATCATCAACGCTCACTACAGGCCCTT 405
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
 DB 406 CACCAAGAGTGGCTTCTCAGACACTCTCAGACAGCAAGAAACAAAGTGAACAACTTCTG 465
 QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
 DB 466 AACCAATGAGCAACCATTTTACAGTACTGGAAGAAAGTTTGCACCCCAAGAACCAAGAA 525
 QY 141 CysLeuLeuGlnGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
 DB 526 TGCTTTATTTACTTGAGCCAGGATTTGATGAATTAATGCGCACAGCAAGCACTACACTCT 585
 QY 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuAspProLeuTyr 180
 DB 586 AGGCTCTGGGCAATGGAGGCTGAGAGGCTGAGGTTGGCAAGCAGCTGAGCCCTGTAT 645
 QY 181 GlnGluTyrValValLeuLysAsnGluMetAlaIleGlnAlaAsnHisTyrGluAspTyrGly 200
 DB 646 GAGAGATATGATGCTGATGTAAGAAACAGATGGCAAGCAACAACTTAAACCATATGGG 705
 QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
 DB 706 GATTTATGAGAGGGGAGACTATGAACAGAGGAGAGAGATGGCTAACACTATACCGCTAAC 765
 QY 221 GlnLeuIleGluAspValGluHisThrPheGlnGluIleLysProLeuTyrGluHisLeu 240
 DB 766 CAGTTGATGATGAGATGATGAACGATCTCCGACAGATCAACCATGATGAGCATCTT 825
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleLys 260
 DB 826 CATGGCTATGATGAGGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 885
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280
 DB 886 TGGCTGCTGCGCCATTTGCTGCTGATATGTTGGTATGTTGGCAAACTGTGACCT 945
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
 DB 946 TTGACTGTTCCCTTGCACAGAAACCAAAATAGATGTTACTGATGATGATGATGATGAT 1005
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320
 DB 1006 GCGTGGATGACGAAGAAGATATTTCAAGAGCAGAGAAATCTTTGTTCTGTTGGCCTT 1065
 QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
 DB 1066 CCTCATATGACCAAGAGATTTGCGGCAAACTTATGCTGACAGCCAGCAGATGGCCGG 1125
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
 DB 1126 ABAAGTTGCTGCAACCCCAAGCTTGGGATCTGGGACAGGAGACTTCAGATCAAGATG 1185
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380
 DB 1186 TGTACAAAGGTCACAAATGAGACAACTTTGAGAGCCCATCAAGATGGAGACATCCAA 1245
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyLysPhe 400
 DB 1246 TATGACATGAGCATATGCGCAGGCAACCTTCTGCTTAAAGAAACGAGCAATGAAGGGTTC 1305
 QY 401 HisGluAlaValGlyLysIleMetSerLeuSerAlaIleTrpProLysHisLeuLysSer 420
 DB 1306 CATCAAGCTGTTGGAGAAATCATGCTTCTGAGCTACCCCAAGCATCTGAATTC 1365

661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680


```

Db 601 GATATGGMGNGGNGATATGARGTNAAYGGNGTGAGVATGATATAYSMNGNGN 660
QY 221 GATATGGMGNGGNGATATGARGTNAAYGGNGTGAGVATGATATAYSMNGNGN 660
Db 661 GATATGGMGNGGNGATATGARGTNAAYGGNGTGAGVATGATATAYSMNGNGN 660
QY 241 HsAlaTyrValArgAlaLysLeuMetLysAlaTyrProSerTyrLysLeuMetLys 260
Db 721 CAYGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 261 CysLeuProAlaHisLeuLeuLysMetTyrGlyArgPheTyrThrAsnLeuTyrSer 280
Db 781 TGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 281 LeuThrValProPheGlyGlnLysProAsnLysLeuValThrAspAlaMetValAspGln 300
Db 841 TTNACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 301 AlaTyrPaspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerValGlyLeu 320
Db 901 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 321 ProAsnMetThrGlnGlyPheThrGlnLysSerMetLeuThrAspProGlyAsnValGln 340
Db 961 CCNAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 341 LysAlaValCysHisProThrAlaTyrPaspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1021 AARGCNGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 361 CysThrLysValThrMetAspPheLeuThrAlaHisIleGlnMetGlyHisIleGln 380
Db 1081 TGTACNAARGTNAARGTNAARGTNAARGTNAARGTNAARGTNAARGTNAARGTNA 1140
QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyPhe 400
Db 1141 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 401 HisGlnAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
Db 1201 CAYGARGCNGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 421 IleGlyLeuLeuSerProAspPheGlnGlnAspAsnGlnIleThrGlyLysPheLeu 440
Db 1261 ATGCGTNTTNTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTyrArg 460
Db 1321 AARGCNGTNTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 461 ThrMetValPheLysGlyGluIleProLysAspGlnIleThrMetLysLysTyrPheGln 480
Db 1381 TGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 481 LysArgGlnLeuValGlyValGlnProValProHisAspGlnIleThrTyrCysAspPro 500
Db 1441 AARGCNGTNTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520
Db 1501 GCNNSTNTTNTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 521 TyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHis 540
Db 1561 TATCATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1620
QY 541 LysCysAspIleSerAsnSerThrGlnAlaGlnGlnLysLeuPheHisMetLeuArgLeu 560
Db 1621 AARTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1680
QY 561 GlyLysSerGlnProThrPheLeuAlaLeuGlnAsnValValGlyAlaLysAsnMetAsn 580

```

```

Db 1681 GGNAAARMSNGARCCNTGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
QY 581 ValArgProLeuLeuAsnTyrPheGlnProLeuPheThrThrPheLysAspGlnAsnLys 600
Db 1741 GTNMGCCNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNT 1800
QY 601 AsnSerPheValGlyTyrPheSerThrAspTyrSerProTyrAlaAspGlnSerIleVal 620
Db 1801 AAYSMNTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGlnIleThrPaspAsnAspGlnMet 640
Db 1861 MGNTTHTSNTTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNT 1920
QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnIleThrPheLeuLysValLysAsn 660
Db 1921 TATYNTTNTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
QY 661 GlnMetIleLeuPheGlyGlnGlnAlaPaspValArgValAlaAsnLeuLysProArgIleSer 680
Db 1981 CATATGCTHTTNTTNTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGlnVal 700
Db 2041 TTYAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
QY 701 GlnLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2101 GARAARCCNTHTGNAATGNAATGNAATGNAATGNAATGNAATGNAATGNAAT 2160
QY 721 SerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2161 MSNTTNAKRTTNTTNTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
QY 741 IleThrLeuIleValPheGlyValAlaMetGlyValIleValValGlyIleValIleLeu 760
Db 2221 ATHTGCTTNTHTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780
Db 2281 ATTTTACNGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAspProGlyPheGlnAsnThrAspAsp 800
Db 2341 TAYGCMWSMNTHTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
QY 801 ValGlnThrSerPhe 805
Db 2401 GTNCCARACNMSNTTY 2415

```

RESULT 14
 ID ABK30270/c
 ID ABK30270 standard, cDNA, 2350 BP.
 ABK30270;
 23-APR-2002 (first entry)
 Human G-protein-coupled protease #40.
 Human; ss; gene; G-protein-coupled protease; gene therapy;
 transgenic; protease mediated disorder; proliferative disorder;
 kw differentiative disorder; developmental disorder;
 kw haematopoietic disorder;
 OS Homo sapiens.
 PN US631427-B1.
 XX 18-DEC-2001.
 PD 26-MAR-1999; 99US-0280116.
 XX
 XX

PR 26-MAR-1999; 9905-0280116.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA
 XX
 PT Robison KE;
 XX
 DR WPI; 2002-129545/17.
 XX
 PT New polynucleotides encoding protease homologs of the G-protein-coupled
 XX protease family, useful in identifying agonists and antagonists for
 XX diagnosis and treatment of protease mediated disorders -
 PS Disclosure; Column 91-94; 246pp; English.
 XX
 CC The invention relates to an isolated human protease nucleic acid molecule
 CC comprising a nucleotide sequence of 546 base pairs, one of 268 fully
 CC defined in the specification. Also disclosed are production of an
 CC isolated polypeptide encoded by the nucleic acid, comprising introducing
 CC the nucleic acid into a host cell and culturing under conditions to
 CC express the protein from the nucleic acid, use of an antibody to
 CC detect the encoded protein in a sample and to modulate its in vivo
 CC activity, identifying agents that bind to the protein and identification
 CC of a polynucleotide agent that modulates the expression of the nucleic
 CC acid or its complement (i.e. gene therapy). The nucleic acid can be used
 CC to identify an agent that modulates the expression or activity of the
 CC nucleic acid, and can be used to isolate the protein. The nucleic acid
 CC can be used in diagnostic assays for determining nucleic acid expression
 CC as well as activity in the context of a biological sample (e.g., blood,
 CC serum, cells, tissue) to determine whether an individual has a disease or
 CC disorder, or is at risk of developing a disease or disorder, associated
 CC with aberrant expression or activity of the nucleic acid. The nucleic
 CC acid can be used to detect mutations in protease genes and gene
 CC expression products such as mRNA. The nucleic acid can be used as
 CC a hybridisation probe to detect naturally-occurring genetic mutations in
 CC a protease gene. The nucleic acid can be used in drug screening methods
 CC to identify agonists and antagonists that can be used to diagnose and
 CC treat such protease mediated disorders e.g., proliferative,
 CC differentiative, developmental or haematopoietic disorders. The nucleic
 CC acid can be used as probes, primers, in biological assays, to determine
 CC patterns of gene expression, to design ribozymes and to construct
 CC transgenic animals. The present sequence represents one of the 268
 CC disclosed human G-protein-coupled protease cDNA sequences.
 XX
 XX
 SQ Sequence 2350 BP; 603 A; 548 C; 489 G; 710 T; 0 other;
 Alignment Scores:
 Pred. No.: 7.28e-311 Length: 2350
 Score: 3291.00 Matches: 726
 Percent Similarity: 92.62% Conservative: 2
 Best Local Similarity: 92.37% Mismatches: 12
 Query Match: 76.70% Indels: 46
 DB: 24 Gaps: 2
 US-09-978-385-2 (1-805) x ABR30270 (1-2350)
 QY 1 MetSerSerSerSerTirpleuLeuSerLeuValAlaVal-ThrAlaAlaGlnSerTh 20
 DB 2305 AGTGAAGCTTCTGCTGCTCTCTCAAGCTTGTCTGTAGATGCTGCTCAATCCAC 2246
 QY 20 rllgluglgluAlaValThr-PheLeuAspLysPheAsnHisGlu-AlaGlnAspLeu 39
 DB 2245 CATTGGAGAACAGCCAAACATTTTGGACAAAGTTTAAACACCAAGAGAGAGCTG 2186
 QY 40 PheTyrGlnSerSerLeuAlaSerTirpaSnTyrAsnThr-AsnIleThrGluLysAsnVa 59
 DB 2185 TTCTATCAAGTTCACCTCTCTTGTGAATTTAAACCCATATTACTGAAGAAGAAATGT 2126
 QY 59 LglAsnMetAsnAlaGlyAsp-LysTirpSer-AlaPheLeuLysGlnSerThr 78
 DB 2125 CCAAAACATGAATATATGCTGGGACAAATGCTGCTTTTAAAGAACAGTCCACA 2066
 QY 79 LeuAlaGlnMetLysTirpLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 98
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 2065 CTTCGCCAAATGTATCATCACTACAAAGAAATTCAGAAATCTCACAGTCAACCTTCATCTGCAG 2006
 QY 99 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 118
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 2005 GCTCTTCAACAAAATATGGGCTTCTCACTGCTCTCAGAAAGACAAAGAGCAAGGCTTGAACACA 1946
 QY 119 IleLeuAsnThrMetSerThrIle-TyrSerThrGly-LysValCysAsnProAspAsn 138
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1945 ATTCTTAATATCAATAGACCACTCATACAGTCTGAGAAAAGTTTGTATACCAAGATATATC 1886
 QY 138 roGlnGluCysLeuLeuLeuGlnLupProGlyLeuAsnGluIleMetAlaAsnSerLeuAsp 158
 DB 1885 CACAAGAAATGCTTATTTACTTGAACACAGCTTGAATGAATATGCAAAACAGTTTAACT 1826
 QY 158 YrAsnGluArgLeuTirpaIleTirpGlnSerTirpArgSerGluValGlyLysGlnLeuArg 178
 DB 1825 ACATATAGAGGCTCTGGGCTTGGGAAAGTGGAGATCTGAGGTCGGCAAGAGCTGAGGC 1766
 QY 178 roLeuTyrGlnGluLysTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGlu 198
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1765 CATTTATATGAAGATATGTGTGCTTGAATAAATGAGATGGCAAGACCAATCATTTAGAG 1706
 QY 198 sPTyrGlyAspTyrTirpArgGlyAspTyrGlu-ValAsnGlyVal-AspGlyTyrAspTyr 217
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1705 ACTATGGGATATTTGGAGAGAGACTATGAATAAAGGGGTAAAGTCTATGACTA 1646
 QY 217 r-SerArgGlyGlnLeuIleGluAspValGlnHisThr-PheGlnGluIleLysProLeu 236
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1645 CATGCCGCGGCGAGTGTGATGAAGATGTGGAACATATCTTTGAAGAGATTAACCATTA 1586
 QY 237 TyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIle 256
 DB 1585 TATGAACATCTTCATCTCATGTGAGGCGCAAGTTGATAAAGCCATTCCTTCATATAC 1526
 QY 257 SerProIleGlyCysLeuProAlaHisLeu-LeuGlyAspMetTyrGly-ArgPheTyr- 275
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1525 AGTCCATTTGATGATCCCTCCCTGCTCATTTCCGTTGGTATATGTGGGCTTAATTTGGG 1466
 QY 276 ThrAsnLeuTyrSerLeuThrVal-ProPheGlyGln-LysProAsnIle-AspValThr 294
 DB 1465 ACAATCTGTACTTGTGACAGATGCTCTTGTGACGGAACCAACATACATCATGTACT 1406
 QY 295 AspAla-MetValAspGlnAlaTirpAspAla-GlnArgIlePheLysGluAla-GlyLys 313
 DB 1405 GATGCCAATGTGTGACAGGCTGTGGATGTGACAGGAATATTCAGAGGAGCCCGCAAG 1346
 QY 314 PhePheValSerValGlyLeuProAsnMetThrGlnGlyPheTyr-GluAsnSerMetLe 333
 DB 1345 TTCTTTGTATCTGTGTGCTTCTCTAATATGACATCAAGATTTGGGAAATTCATGCT 1286
 QY 333 utThr-AspProGlyAsnValGlnLysAlaValCysHisProThrAlaTirpAspLeu-Gly 352
 DB 1285 AACGGAGCCAGGAATGTTCAGAAAGCATGCTCCATCCACAGCTTGGAGCTGGGGG 1226
 QY 353 LysGlyAsp-PheArgIle-LeuMetCysThrLysValThrMetAspPheLeuThr 372
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1225 AAGGGCCAGTTTCAAGATCTCTTAATGTGCACAAAGGTGACATGACGCTTCTACAG 1166
 QY 372 IaHisHisGluMetCylHisIleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeu 392
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1165 CTGATCATGAGATGGGCAATATCCAGATGATGTGATGTGTCGCAACCTTTCTCTGG 1106
 QY 392 euarqasnglyAlaAsnGluGlyPheHisGlnAlaValGlyGluIleMetSerLeuSer 412
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1105 TAAAGAAATGAGCTTAAAGAGATTCATGACCTGTGGGAAATCATGTCACTTTCTG 1046
 QY 412 IaAlaIlePro-LysHisLeuLysSerIleGlyLeuLeu--SerProAspPheGlnGlu 430
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1045 CAGCCACACCTTAAGCATTTAAATCCATTTGTTTGTACCCGAAATTTTCAAGA 986
 QY 431 AspAsnGluThrGluIleAsnPheLeuLysGln-AlaLeuThrIleValGlyThr-L 450
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 985 GACAATGAACAGAAATAACTTCTCTCAACAAAGACATCAGATGTGTGGACTCC 926

QY 40 PheTyrGlnSerSer-LeuAlaSer-TyrAsnTyrAsnThrAsnIlethrgInGluAsnV 59
 |||||
 Db 159 TTTCTATCAAAAGTTCAAGTTCCTCTGGAAATATATACACCAATATATACAGAGATG 218
 QY 59 aGln-AsnMetAsn---AsnAlaGlyAspLys-----TyrSerAlaPheLeuLysGlu 75
 |||||
 Db 219 TCCAAACAACATGCATTAAGTTGCTGGCAGAACAAATGTGTCTAGCCCTTTTACAGSAA 278
 QY 76 GlnSerThrLeuAlaGln-MetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLe 95
 |||||
 Db 279 CAGTCCACACTTGGCCCAAGATGTATCCACTACAA-----GCCAACTTCACGACACTGCC 332
 QY 95 uGlnLeuGlnAlaLeuGlnGlnAsnGly-----SerSerVal 107
 |||||
 Db 333 ACATGTCAACGCTTCAGCTGCTCAGCGCTTCTTCAAGCCATAAATCTGTAGCTCTTCAG 392
 QY 107 IleuSer-----GluAspLysSerLys-ArgLeuAsnThrIleLeuAsnThrMet---- 123
 |||||
 Db 393 TTGTTCATCAGACGACACAGAGAGAACCCGGTTGACACAAATTTCTATATACAAATGG 452
 QY 124 -----SerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlnCysL 142
 |||||
 Db 453 AGCCACCAATCTTAACAGTAACTGAGAAACGTCGTACCCAGATATATCCACAGATGCT 512
 QY 142 eulLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGlnArgL 162
 |||||
 Db 513 TATTTACTGACACGAGTTTGAATGAATATGCAACAGTTTATAGACTACATAGAGAGGC 572
 QY 162 eutThrAlaThrGlnSerThrPargSerGluValGlyLysGlnLeuArgProLeuTyrGlnG 182
 |||||
 Db 573 TCTGGGCTTGGGAAAGCTGGAGATCTGAGTGGCAGACGCTGAGCCATTATATGAGAG 632
 QY 182 IuTyrValValLeuLysAsnGlnMetAlaArgAlaAsnIleTyrGlnLysP-Tyr-GlyAs 201
 |||||
 Db 633 AGTATGTGGTCTTGAATAATGAGATGGCAGACCAATATATATAGAGACTATTTGGGGA 692
 QY 201 pTyrTPargGlyAspTyrGluValAsnGlyValAsp--GlyTyrAspTyrSerArgG 220
 |||||
 Db 693 TATTTGAGAGAGACTATGAGTAAATGGGTAAATAGTGATATGATTAACGCCGCGG 752
 QY 220 yGlnLeuIleGlnAspValGlnHisThr-PheGlnGlnIleLysProLeuTyr-GlnHis 239
 |||||
 Db 753 CCAGTTGATTTGAAGATGTGAACATACCTGTTGAAGAGATTAACCATTTGATGAGAACAT 812
 QY 240 LeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIle 259
 |||||
 Db 813 CTTCACGCCCTATGTGAGGCCCAAGTTGATGATGCTTATCTCTATATCTGCATCT 872
 QY 260 GlyCysLeuProAlaHisLeuLeuGlnLysAspMetTyr-GlyArgPheTyrThrAsnLeuTyr 279
 |||||
 Db 873 GGAATGCCCTCCCTGCTCATTTGGTGGTATGTGCGGGTAGATTTTGGACAAATCTGTA 932
 QY 279 rSerLeuThrValProPheGlnGlnLysProAsnIleAspValThrAspAlaMetValAs 299
 |||||
 Db 933 CTCTTTGACAGTTCCTTTGGACAGAACCAACATAGATGTACTGATGACATGGGTGA 992
 QY 299 pGlnAlaTyrPaspAlaGlnArgIlePheLysGlnAla-GluLysPhePheValSerValG 319
 |||||
 Db 993 CCAGGCGCTGGATGACACAGATATATTCAGAGAGTCCGAGAACTCTTTGATCTGCTGG 1052
 QY 319 lLeuProAsnMetThrGlnGlnLysPheTyr-GluAsnSerMetLeu-ThrAspProGlyAs 338
 |||||
 Db 1053 GTCTTCCTTATATAGACTATAGATTTCTGCGGCAAAATTTCCATGCTATACGAGCCAGAAA 1112
 QY 338 nValGlnLysAlaValCysHis-ProThrAlaTyrPaspLeuGlnLysGlyAspPheArg- 357
 |||||
 Db 1113 TGTTCAGAAAGCACTCTGCCATCCCAACAGCTTGGGAGACCTGGGAGAGGCGACTTCAGAG 1172
 QY 358 lIleuMetCysThrLys-ValThrMetAspAspPheLeuThrAlaHisGlnMetAl 377
 |||||
 Db 1173 ATCTTATGTGACAAAGGATACAAATGACGACTTCTCTGACAGCTCATCATGAGATGGG 1232
 QY 377 yHisIleGlnTyrAspMetAlaTyrAlaAlaGlnPro-PheLeuLeuArg-AsnGlyAla 396
 |||||

Db 1233 GCATATCCACTATGATATGCGCATATGCGGCCCAACCTTTTCTTGCTAAGAAATGAGGCT 1292
 QY 397 -AsnGlnGlyPheHisGlnAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLys 416
 |||||
 Db 1293 TAAATGAAGATTCATGCAATGACCTGTGGGAAATCATGTCATCTTTCGACGCCACACTTA 1352
 QY 416 SHSLeuLysSerIleGlyLeuLeuSerProAspPheGln---GluAspAsnGlnThrG 435
 |||||
 Db 1353 GCATTTAAATTCATTTGCTTCTGTCTGACCCGAGTTTTCACGACAGACATGAACACAA 1412
 QY 435 uIleAsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMet 455
 |||||
 Db 1413 AATTAACCTCTGCTCAAAACAGACACATGATTTGGGAGCTGACCTGCAATTTACTTACAT 1472
 QY 455 tLeuGlnLysTyrPargTyrPheThrValPheLys-GlyGlnIleProLysAspGlnTyrMetL 475
 |||||
 Db 1473 GTTAGAAGATGGAGGTGGATGGCTTTAAAGCGGGAAATTCACAAAGACCAAGTGGGTGA 1532
 QY 475 ySLys-TyrTPargIleMetLysArgGlu-IleValGlyVal-ValGluProValProHisA 494
 |||||
 Db 1533 AAAAGTGTGGGAGATGAGAACCGAAGAAATAGTTGGGTGTGTGGAACCTGTGCCCATG 1592
 QY 494 srgIuThrTyr-CysAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIle 513
 |||||
 Db 1593 ATGAACATATCTGTGACCCCGCATCTCTGCTCCATGTTTCTAATGATTAATCTATTCAT 1652
 QY 514 ArgTyrTyrThrArgThrLeu-TyrGlnPheGlnPheGln-GluAlaLeu-CysGlnAla 532
 |||||
 Db 1653 CGATTTATACACAGAGACCCGTTTACCAATTCAGTTTCAAAAGAGACCTTTGTCAAGCA 1712
 QY 533 AlaLysHisGlnGlyProLeuHisLys-CysAspIle-SerAsnSerThrGlu--AlaG 551
 |||||
 Db 1713 GCTAAACATGAGGCGCTTCGACAAATGTGACATTCCTCAAAATTCACAGAAAGCTGCTG 1772
 QY 551 yGlnLys-LeuPheAsnMetLeuArgLeuGlyLys-SerGluProThrPheThrAlaLe 570
 |||||
 Db 1773 GACAGAACCTGTTCAATATGCTGAGGCTTGGAAACTCGAAGACCTGGACCTTAGCAT 1832
 QY 570 uGlnAsnValVal-GlyAlaLysAsnMetAsnValArgPro-LeuLeuAsnTyrPheGlu 589
 |||||
 Db 1833 GGAAATGTTTGTAAAGACCAAGAACATGATGAAGCCACCTGCTCAACTCTTTGAG 1892
 QY 590 ProLeuPheThrTyrPheLysAspGlnAsnLysAsnSerPheValGlyTyrPserThrAsp 609
 |||||
 Db 1893 CCTTATTTACTGCTGGAAGAACCAAGAACATCTTTTGGGATGAGATACCGAC 1952
 QY 610 TrpSerProTyrAlaAspGlnSerIle-LysValArgIleSerLeuLysSerAlaLeuG 629
 |||||
 Db 1953 TGGAGTCCATATGACAGACACAGCATCACAAGTGAAGATGAAGCTTAATAACAGCTTGG 2012
 QY 629 y-AspLysAlaTyrGluTyrAsnAsp-AsnGlnMetTyrLeuPheArgSerSer-ValAl 648
 |||||
 Db 2013 CAGATTAAGCATATGATGAGACGACCAATGAATGATACCTGTTCCGATCATCTGGTTGG 2072
 QY 648 aTyrAla--MetArgGlnTyrPheLeu-LysValLysAsnGlnMetIleLeuPheGlyG 667
 |||||
 Db 2073 ATATTGTAATTTGAGGAGTACTTTTAAACAAGTAAATAATCGATATCTTTTGGGG 2132
 QY 667 lGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePheValThrA 687
 |||||
 Db 2133 AGGAGGATGTGCGAGTGGCTAATTTGAAACCAAAATCTCTTTAATTTCTTTGTCACTG 2192
 QY 687 lAspLysAsnValSer-AspIleIleProArg-ThrGluValGluLysAlaIleArgMe 706
 |||||
 Db 2193 CACCTTAATAAGTGTGCTGATATCATCTCTTGAACACGAGTTGAAGGACCATCAGAT 2252
 QY 706 tSerArgSerArg-IleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGlnPheLeuG 726
 |||||
 Db 2253 GTCCCGGAGCGCTACTCATGATGCTTCCCTGTAATGACAGACAGCTTGAAGTTTCTGG 2312
 QY 726 lYlIleGlnProThrLeuGlyProProAsnGlnProProValSerIleThrLeuIleValP 746
 |||||

	CC	conditions including left ventricular systolic dysfunction, progressive
	CC	renal impairment, scleroderma renal crisis, congestive heart failure due
	CC	to dysfunction, and treatment of atherosclerosis. Zacc2 agonists may be
	CC	used to treat infertility while Zacc2 antagonists are used for inducing
	CC	infertility. The present sequence represents a degenerate sequence
	CC	encoding the mouse Zacc2-5 protein.
	XX	
50	Sequence	2415 BP; 488 A; 228 C; 397 G; 324 T; 978 other:
	Alignment Scores:	
	Pred. No.:	1,36e-272 Length: 2415
	Score:	2899.00 Matches: 526
	Percent Similarity:	71.06% Conservative: 46
	Best Local Similarity:	67.56% Mismatches: 233
	Query Match:	67.56% Indels: 0
	DB:	22 Gaps: 0
US-09-978-	-385-2 (1-805) x AAC84369 (1-2415)	
OY	1 MetSerSerSerSerTrrPleuLeuLeuSerLeuValAlaValThrAlaAagInSerThr	20
Db	1 ATGSMNSMNSMNSMNSTGGTYNTYTNTWMSYNTGTNCGCTNACNGCNGCMWSNYTN	60
OY	21 lIeGluglInalAlaLyStrPhelEuAsPlysPheasnIgLuAgluaSplEuPhe	40
Db	61 ACNGARCARAAyGCNNARACNTTYTTAAATAATTAAyCARBARCGNCARGAYTNMWSN	120
OY	41 TyrcInSerSerLeuAlaSerTTPasrTyTAstThraSnIlEthrGlugluAsnValGln	60
Db	121 TATCAARMSMNSYNTGCMWSMTGAATATAAAACAANAyTHACNGARCARAAyGCNCAR	180
OY	61 AsmEtasnaSnAlaGlyAsPlysTPSPSeAlaPheLeuSgltugInserThrLeuAla	80
Db	181 AARTGTGSNARCGNCGNCNAARTGTGMSGCTTTTATGARBARAKARMSNARACNGCN	240
OY	81 GlmMetTyrrProLengIngLuIlleglaSnLeuThrVallysLeuIngLInalAleu	100
Db	241 CARMSNTTYSNTNCARGARATHCARACNCCNATATHAARMGNCARFTNCARGCNTYN	300
OY	101 GlnglInaSnGlySerSerValLleuSerGIuAsPlysSerLyArgLeuasnThrIleLeu	120
Db	301 CARCARMSNCGMNSMNSGCCNTYTNMSNGCAGATAAAATAARCAARTNAyVCMATHHTN	360
OY	121 AsnThrMeSerThrllEtyrSerThrGlyLyValCyAsnPrOksasnProGlnglu	140
Db	361 AAACANATGSMNACNATNTFAyMSNACNGMAANGTWTGYAACCMNARAyCCNCRGAR	420
OY	141 CySleuLeuLeuEngLurProGlyLeuAsnGluIlleMetAlaAsnSerLeuAspTyraSngu	160
Db	421 TGTYNTYNTNTNARCRCNGNTNYNGAYGARATHAGCGCACMWSMNACNAYTAyVAyWSN	480
OY	161 ArgLeuTrpAlarProLusErTPArySerSerGIuValGlyGIuLeuArgProLeuTyrr	180
Db	481 MGyVTMGCGCINTGGARGRGNTGCMONGCNGARGTNGMAACARTTMNGNCNTTNAy	540
OY	181 GluGluTyValValLeuLysasnGluMetAlaArGAlaAsnHisTyrcIuAspTyrgly	200
Db	541 GAyGARTAyGTCNTNTNARAyGARARAGCGNMGCNAAAATAyTAyAAyATyAGN	600
OY	201 AspTyrrTriPArGLuAspyrGIuValasnglyValasglyValasglyAspTyrrargly	220
Db	601 GATTAyTGGMGNNGNARyTAyGARCGNKGRRGCGCGATGGTAAATAyTAyAAGNAY	660
OY	221 GluLeuIleGluAspValGluHisThrPheGluGluIlleLySproLeuTyrcIuHisLeu	240
Db	661 CAyRTTATATGARGAyTCtNARBMCNCTTYGCNGARATHAARCNYTWTAYGARCAyTYN	720
OY	241 HisAlaTyrrValArgAlaLysLeuMetAsnAlaTyrrProSerTyrrIleSerProIlleGly	260
Db	721 CATGCATATGTMGNMGNAARTTNATGGAYACNTAYCCMWSMTAATATMSNCCNACNGN	780
OY	261 CySleuProAlaHisLeuLeuLysaspMetTyrrGlyTyrrPheTyrrThrAsnLeuTyrrSer	280

DR P-PSDB; AAR10426.

XX Nucleic acid - encoding human testicular angiotensin conversion
 PT enzyme, used e.g. for in vitro detection of enzyme in organism
 XX
 PS Claim 1; Fig 1; 48pp; French.

XX
 CC A bank of human testicular cDNA in Lambda gt11 was screened with a
 CC probe containing the final 3248 nucleotides of endothelial ACE. The
 CC complete sequence of ACE was reconstructed from 4 separate clones.
 CC It encodes a 711 amino acid mature protein and a 21 residue signal
 CC peptide. The 228-2224 sequence is identical to the 1944-3940 region
 CC of endothelial ACE. The isolated nucleic acid sequence is inserted
 CC into a plasmid for expression of polypeptides. The invention also
 CC covers parts of the sequence comprising all or part of the 29-229
 CC sequence, any sequence differing from ACE only by silent
 CC substitutions and nucleic acids which hybridise to ACE.

XX
 SQ Sequence 2477 BP; 536 A; 811 C; 695 G; 435 T; 0 other;

Alignment Scores:

Pred. No.:	8-01e-121	Length:	2477
Score:	1344.00	Matches:	259
Percent Similarity:	60.97%	Conservative:	119
Best Local Similarity:	41.77%	Mismatches:	204
Query Match:	31.32%	Indels:	38
DB:	12	Gaps:	10

US-09-978-385-2 (1-805) x AAQ10328 (1-2477)

OY 15 ThrAlaAlaGlnSer-----ThrIleGlnGlnAlaAlaGlnSerPheLeuAsp 30
 DB 209 ACATAGCGCCAGACCCAAACCTGCTGACTGATGAGCGTGAAGCCAGCAAGTTTGAGAG 268
 OY 31 LysPheAsnHISGluAlaGluAspLeuPhePheGlnSerSerLeuAlaSerTrpAsnTyr 50
 DB 269 GAATATGACCGGACATCCAGAGGTGTGTGAAGAGATGTCGCGAGCACTGTGAACATC 328
 OY 51 AsnThrAsnIleThrGluGlu-----AsnValGlnAsnMet 62
 DB 329 AACACCAACATCACCACAGACAGCAAGATTCGTCGACAGAGAACATCAATATGACC 388
 OY 63 AsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAlaGlnMet 82
 DB 369 AACCCACACC-----CTGAAGTACGGCACCCAGCCAGGACAGAG 424
 OY 83 TyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnIleGlnAlaLeuGln 102
 DB 425 TTTGATGTGAACCAAGTTCAGAACACCACTATCAAGCGGATCATTAAGAGGTTCCAGAGC 484
 OY 103 AsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeuAsnThr 122
 DB 485 CTAGAACGGGAGCGCTCCTGCTGCCAGAGCTGGAGGAGTCAACAGATCTCTTGAT 544
 OY 123 MetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnProGlnGlnCysLeu 142
 DB 545 ATGGAACACCACTACAGCGTGGCCACTGTGTGCCACCCGAGATGCC-----AGCTGCTG 598
 OY 143 LeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGluArgLeu 162
 DB 599 CAGCTCGAGCCAGATCTGAGCAATGTATGTGCCACATCCCGAATATGAAGACCTCTTA 658
 OY 163 TrpAlaTrpGlnSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyrGluGlu 182
 DB 659 TGGGATGTGGAGGCTGGCAGACAGCGGGAGACCACTCTCCAGTTTACCCGAAA 718
 OY 183 TyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGlnAspTyrGlyAspTyr 202
 DB 719 TACGTGAACATCAACACAGGCTGCCCGCTCAATGCTATGTATGACAGGGGACTCG 778
 OY 203 TrpArgLysAspTyrGlnValAsnGlyValAspGlyTyrAspTyrSerArgGlyGlnLeu 222
 DB 779 TGGAGGTCTATGTACGAGACCACTCCCTGGAG----- 811

OY 223 IleGluAspValGlnHisThrPheGlnGluIleLysProLeuTyrGlnHisLeuHisAla 242
 DB 812 ---CAAGACCTGGAGGCGCTCTCCAGAGCTGCAGACCATCTCAACCTGCATGCC 868
 OY 243 TyrValArgAlaLysLeuMetAsnAlaTyr---ProSerTyrIleSerProIleGlyCys 261
 DB 869 TACGTGGCCCGGCGCTGCACCTGACTACAGGGGCCAGACATCAACCTGAGGGGCC 928
 OY 262 LeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpHisLeuTyrSerLeu 281
 DB 929 ATTCTGTCTACCTGCTGGAGACATGTGGGCGACCTGGTCAACATCTATGACTGT 988
 OY 282 ThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGlnAla 301
 DB 989 GTGGTGGCTTCCTTCAGGCCCTTCATGAGCCACACAGAGGTATGCTAAAGCAGCGC 1048
 OY 302 TrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeuPro 321
 DB 1049 TGGAGCCCGCAGAGAGATGTTAAAGAGCTGATGATTTCTTCACTCCCGGGGCTGCTG 1108
 OY 322 AsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGlnLys 341
 DB 1109 CCGGTGCTCTCTGATGCTTGTGACACATGCTGAGAGAGCCACCGAGGGGCGGAG 1168
 OY 342 AlaValLysHisProThrAlaTrpAspLeuGlyLysGly---AspPheArgIleLeuMet 360
 DB 1169 GTGGTGTGCACAGCGCTGGCGCTGGAGCTGTACAAAGCAAGAGCTCCGATCAAGCAG 1228
 OY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGlnMetGlyHisIleGln 380
 DB 1229 TGCACCACTGAACTGGAGAGACCTGTGTGTGCCCCACCAAAATGGGCCCATATCCAG 1288
 OY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPhe 400
 DB 1289 TATTTCATGCAGTAACAAGACTTACCTGTGGCTTGAGGAGGAGGTGCCAACCCGGCTTC 1348
 OY 401 HisGlnAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 DB 1349 CATGAGGCAATGTGGGAGCGGTGACGCTCTCAGTGTACAGCCCAAGCACTGCACAGT 1408
 OY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGlnIleAsnPheLeu 440
 DB 1409 CTCAACTGCTGAGAGAGTAGGGTGGCAGAGCAG---GAGCATGACATCAACTTTCGATG 1465
 OY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTrpArg 460
 DB 1466 AAGATGCCCCCTTGACAAGATCGCCCTTATCCCTTGACGCTACCTCGATCAGTGGCGC 1525
 OY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpArgLeuMet 480
 DB 1526 TGGAGGTTATTTATGGAACATCAACAGAGAACTATTAACAGGAGGTGTGGAGCCCTC 1585
 OY 481 LysArgGluIleValGlyValValGluProValProHisAspGluThrTyrGlyAspPro 500
 DB 1586 AGGCTGAAGTACAGAGGCGCTGCCCCAGAGTCCAGAGACAGTCAAGCTTATGACCA 1645
 OY 501 AlaSerLeuPheHisLysSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520
 DB 1646 GGGGCCAAGTTCACATCTCTTCTAGCGTCCCTTACATAGGTAACCTTGTGACCTTCATC 1705
 OY 521 TyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGlnLysProLeuHis 540
 DB 1706 ATCACTGTCACAGTCCACAGAGAGCTGTGCCAGACAGCTGGCCACAGGGGCCCTGACC 1765
 OY 541 LysCysAspLysSerAsnSerThrGlnAlaGlyLysLeuPheAsnMetLeuArgLeu 560
 DB 1766 AAGTGTGACATATCACAGTCCAAAGAGCGGGGAGCGCTGGCGAGCCCATGAAGCTG 1825
 OY 561 GlyLysSerGlnProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetLys 580
 DB 1826 GGTCTTGAAGAGCGCTGGCGGAAAGCCATGATCAAGGCGGAGCCCAACATGAGC 1885

QY 581 ValArgProLeuLeuAsnTyr-PheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600
 Db 1886 GCCTGGCCCATGTTAGCTACTTCAAGCCGCTGCTGAGCTGGCTCCGACGAGACAG 1945
 QY 601 -----AsnSerPheValGlyTrp---SerThrAspTrpSerProTyrAlaAspGlnSer 617
 Db 1946 CTGCATGGGAGAACACTGGGCGGACGAGTACACTGACGCGGACACTCGCTGCTCA 2005
 RESUIT 18
 AAA38330
 ID AAA38330 standard; DNA; 4020 BP.
 AC AAA38330;
 XX
 XX 21-AUG-2000 (first entry)
 DT
 XX Human angiotensin-converting enzyme (ACE) coding region.
 DE
 XX Angiotensin-converting enzyme gene; ACE; coding region; polymorphism;
 KW polymorphic marker; cardiovascular disease; myocardial infarction;
 KW unstable angina; hypertension; atherosclerosis; stroke; prognosis;
 KW drug screening; treatment outcome; human; ds.
 XX
 XX Homo sapiens.
 OS
 XX WO200022166-A2.
 PN
 XX 20-APR-2000.
 PD
 XX 13-OCT-1999; 99WO-IB01678.
 PF
 XX 14-OCT-1998; 98US-0104286.
 PR 14-OCT-1998; 98US-0104302.
 XX
 XX (EURO-) EUROPA MEDICAL AB.
 PA
 XX Norberg LT, Andersson MK, Lindstrom PHR, Jonsson L;
 PI
 XX WPI: 2000-318010/27.
 DR
 XX
 XX Assessing cardiovascular status in humans involves comparing test
 PT polymorphic pattern comprising polymorphic positions within genes
 XX encoding specific proteins, with reference polymorphic pattern
 XX
 PS Disclosure: Page 114-115; 126pp; English.

CC treatment regimen. Adverse results in an early trial can be evaluated to
 CC identify polymorphic patterns so that the adverse results can be
 CC correlated with a sub-population of the test population, permitting
 CC exclusion of such sub-populations from the treatment group. Beneficial
 CC drugs can be approved for use in the appropriate population, thereby
 CC decreasing the number of patients required for a clinical trial, which in
 CC turn decreases the duration and cost of such trials. Sequences A38328 and
 CC A38330 represent, respectively, intron 16 and the coding region of
 CC the human ACE gene (GenBank X52855, J04144). The polymorphic sites
 CC identified are 375A/C, 582C/T, 731A/G, 1060G/A, 1215C/T, 2193G/A,
 CC 2328A/G, 2741G/T, 3132C/T, 3387T/C, 3503G/C, 3906G/A; and a deletion of
 CC nucleotides 1451-1783 in intron 16.
 CC
 XX
 SQ Sequence 4020 BP; 857 A; 1261 C; 1174 G; 728 T; 0 other;
 Alignment Scores:
 Pred. No. 7.73e-120 Length: 4020
 Score: 1337.00 Matches: 235
 Percent Similarity: 61.05% Conservative: 118
 Best Local Similarity: 41.73% Mismatches: 204
 Query Match: 31.16% Indels: 34
 DB: 21 Gaps: 9
 US-09-978-385-2 (1-805) x AAA38330 (1-4020)
 QY 20 ThrIleGluGluGlnAlaLysTrpPheLeuAspLysPheAsnHisGluAlaGluAspLeu 39
 Db 1952 ACTGATGAGCGGTGAGCGACACAGTTGTGTGAGGAATATGACCGGACATGCCAGTGGTG 2011
 QY 40 PheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGlu----- 57
 Db 2012 TCGAACAGATATGCCGAGCACTGCACTACACACCCACATCCACACAGACAGCAGC 2071
 QY 58 -----AsnValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAla 71
 Db 2072 AAGATTCTGCTGCAGAAACATGCAATATGCCACACACAC----- 2113
 QY 72 PheLeuLysGluGlnSerTrpLeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeu 91
 Db 2114 -----CTGAAGTACGGGACCCAGCCAGGAAAGTTGATGTGAACCGTGTGACAGACACC 2167
 QY 92 ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnLysSerValLeuSerGluAsp 111
 Db 2168 ACTATCAAGCGCATATTAAGAAAGTTACAGACCTAGACAGGCGGCTGCTGCCAG 2227
 QY 112 LysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLys 131
 Db 2228 GAGCTGAGAGGATACACACAGATCTGTGTGATATGGAACACCTACAGCGGCCACT 2287
 QY 132 ValCysAsnProAspAsnProGlnGluCysLeuLeuGluLupProGlyLeuAsnGluIle 151
 Db 2288 GTGTGCCACCCGAATGGC-----AGCTGCTGCAGCTGACGACAGATCTGACGAATGTG 2341
 QY 152 MetLysAsnSerLeuAspTyrAsnGluArgLeuTrpAlaTrpGlnSerTrpArgSerGlu 171
 Db 2342 ATGGCCACATCCCGGAATATGAAGACCTGTATTGGCCATGGAGGCTGGCAGACAG 2401
 QY 172 ValGlyLysGlnLeuArgProLeuTyrGluGluValValLeuLysAsnGluMetAla 191
 Db 2402 GCGGGAGAGGCACTCCATCTTATCCCGAATACCTGGAACATCATACACAGGCTGCC 2461
 QY 192 ArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGly 211
 Db 2462 CGGTCATATGCTATGTAGATGACGAGGAGACCTGCGAGGCTATATGTCGAGACACATCC 2521
 QY 212 ValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGlnHisThrPheGlu 231
 Db 2522 CTGGAG-----CAAGACCTGGAGCGGCTTTCAG 2551
 QY 232 GluIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAla 251
 Db 2552 GAGCTGCAGGCACTCTACCTGACCTGATCTACGTCGCGGCGGCGCTGCACCGTCCAC 2611

QY 252 Tyr---ProSerTyrIleSerProIleGlyCysLeuProAlaHisIleLeuGlyAspMet 270
 Db 2612 TACGGGCCCCAGACACATCATGAGTGGCCCTTCTCCCTGACCTCTGGGGAGACATG 2671
 QY 271 TrpGlyArgPheTrpThrAsnLeuTyrSerLeuThrValIleProPheGlyGlnLysProAsn 290
 Db 2672 TGGGGCCAGACTGGTCCACACATCATGAGTGGCCCTTCTCCCTGACCTCTGGGGAGACATG 2731
 QY 291 IleAspValThrAspAlaMetValAspGlnAlaTrpAspAlaGlnIlePheLysGlu 310
 Db 2732 ATGACACACACAGAGGATCTGTAAGACAGGCGCTGGACCCAGAGAGATCTTTAAAGAG 2791
 QY 311 AlaGlyLysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheTrpGluAsn 330
 Db 2792 GCTGATGATTTCTTCCACTCCCTGCGGCTGCTGCGCTCCCTGAGTCTGGAAACAG 2851
 QY 331 SerMetLeuThrAspProGlyAsnValGlnLysAlaValCysHisProThrAlaTrpAsp 350
 Db 2852 TCGATGCTGGAGAACCCAGACCGGCGGAGGTGGTGTGCTGCACGCTCGGCTGGGAC 2911
 QY 351 LeuGlyLysGly---AspPheArgIleLeuMetCysThrLysValThrMetAspAspPhe 369
 Db 2912 TTCTACACAGCGCAAGACTCTCCGATCAAGCAGCACCAGCTGAGACTTGGAGACTG 2971
 QY 370 LeuThrAlaHisIleGlyMetGlyHisIleGlnTyrAspMetAlaTyrAlaAlaGlnPro 389
 Db 2972 GTGGTGGCCACACGAATGGCCACATCCAGTATTTCAAGTACCAAGACTTACT 3031
 QY 390 PheLeuLeuArgAsnGlyAlaAsnGlyLeuPheHisGlnAlaValGlyGlnIleMetSer 409
 Db 3032 GTGGCTTGAAGAGGAGGCGCCACCCGCGCTTCCATGAGCCATTTGGGAGGTCTAGCC 3091
 QY 410 LeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuSerProAspPheGln 429
 Db 3092 CTCTCAGTCTAGCGCCACACACCTGACAGCTCAACCGCTGAGCAGAGGAGGCGC 3151
 QY 430 GluAspAsnGlnThrGlnIleAsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThr 449
 Db 3152 AGCGAC---GAGCATGACATTAACCTTCTGTGAAGATGGCCCTTGACAAAGATGCTTT 3208
 QY 450 LeuProPheThrTyrMetLeuGlyLysTrpArgTrpMetValPheLysGlyGlnIlePro 469
 Db 3209 ATCCCTTCACCTACCTGCTGATGATGAGTGGCGCTGAGAGGATTTGATGCAAGCATACC 3268
 QY 470 LysAspGlnTrpMetLysLysTrpTrpGlnMetLysArgGlnIleValGlyValGlu 489
 Db 3269 AAGGAGACATATAACACAGAGTGGTGGAGCTGACCTGAAAGTACAGAGGCGCTGCCCC 3328
 QY 490 ProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSerAsnAsp 509
 Db 3329 CCAAGTCCCGACGACTCAAGGATGACTTGGACCCAGGCGCAAGTTCCATCTCTTACG 3388
 QY 510 TyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnGlnAlaLeu 529
 Db 3389 GTGGCTTACATCAGTACTTGTTCAGCTTCATCATCCAGTTCCAGTTCCAGGAGCATG 3448
 QY 530 CysGlnAlaAlaLysHisGlyGlyProLeuHisLysCysAspIleSerAsnSerThrGlu 549
 Db 3449 TGGCAGGACACTGGCCACAGGCGCCCTGACAAAGTGTACATCTACAGTCCCAAGAG 3508
 QY 550 AlaGlyLysLeuPheAsnMetLeuArgLeuGlyLysSerGlnProTrpThrAla 569
 Db 3509 GCGGGGAGGCGCTGGCGACCCGCAATGAAGTGGCTTCAAGTGGCGCGGAGAGCC 3568
 QY 570 LeuGlnAsnValGlyAlaLysAsnMetCysValArgProLeuLeuAsnTyrPheGlu 589
 Db 3569 ATGACGCTGATCAAGGCGCCAGCCCAACATAGAGGCGCTCGGCAATGTTGACTCTTCAAG 3628
 QY 590 ProLeuPheThrTrpLeuLysAspGlnAsnLys-----AsnSerPheValGlyTrp--- 606
 Db 3629 CCGTCTGTGACTGGCTCCGCGACGAGAAAGCATGCTGATGGAGAAAGCTGGCTGGCG 3688
 QY 607 SerThrAspTrpSerProTyrAlaAspGlnSer 617

Db 3689 CAGTACACTGGACGCCGAATCTCGCTGCTCA 3721
 RESULT 19
 AAK93395
 ID AAK93395 standard; DNA; 4022 BP.
 XX
 AC AAK93395;
 XX
 DT 27-JUN-2002 (first entry)
 XX
 DE DNA of APP related human homologue hcp51674.
 XX
 KW Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;
 KW amyloid precursor protein; tissue-specific expression control; human APP;
 KW APP pathway modulator; gene therapy; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 148..3945
 FT /tag= a
 FT /product= "Protein of human homologue hcp51674"
 FT /note= "No start codon"
 XX
 MO200226820-A2.
 XX
 PD 04-APR-2002.
 XX
 PF 01-OCT-2001; 2001MO-BP11345.
 XX
 PR 29-SEP-2000; 2000US-236893P.
 PR 14-JUN-2001; 2001US-298309P.
 XX
 PA (NOVS) NOVARTIS AG.
 PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 XX
 PI Cohen D, Dengler UJ, Finelli AL, Freuler F, Konsolaki M;
 PI Reinhardt MWHM, Zusan S;
 XX
 DR WPI: 2002-315796/35.
 DR P-PSDB; AAO20501.
 XX
 PT New transgenic fly, containing DNA encoding an Abeta portion of human
 PT APP, useful for identifying agents which modulate the APP pathway and
 PT which can be used to treat Alzheimer's disease -
 XX
 PS Example 4; Page 93-94; 129pp; English.
 XX
 CC The invention relates to a transgenic fly whose genome comprises DNA
 CC encoding a polypeptide having the Abeta portion of human amyloid
 CC precursor protein (APP), fused to a signal sequence. The DNA sequence
 CC encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in
 CC the specification. The DNA sequence is operably linked to a tissue-
 CC specific expression control sequence. Expression of the sequence gives
 CC the fly an altered phenotype. The purpose of the invention is for
 CC identifying agents that inhibit or promote the expression and/or function
 CC of genes or encoded polypeptides which modify the APP pathway. The agent
 CC is a compound, triple helix DNA, antisense oligonucleotide, double
 CC stranded RNA molecule, ribozyme, or particularly an antibody. It is used
 CC to treat conditions such as Alzheimer's disease. The agent can be used as
 CC an APP pathway modulator or in gene therapy. This polynucleotide sequence
 CC represents the DNA of the APP related human homologue hcp51674.
 XX
 SQ Sequence 4022 BP; 856 A; 1262 C; 1174 G; 728 T; 2 other:
 Alignment Scores:
 Pred. No.: 7.74e-120 Length: 4022
 Percent: 1337.00 Matches: 255
 Score: 61.05% Conservative: 118
 Best Local Similarity: 41.73% Mismatches: 204
 Query Match: 31.16% Indels: 34
 DB: Gaps: 24

US-09-978-385-2 (1-805) x AAK9395 (1-4022)

QY 20 ThrIleGluGlnAlaIleLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeu 39
 Db 1954 ACTGATGAGGCGTGGAGCCAGCAAGTTTGTGGAGCAATATGACCGGACATCTCCAGGTGTG 2013
 QY 40 PheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGlu----- 57
 Db 2014 TGGAAAGCATATGCGAGGCCACTGGAATACACACCAATCACCACAGACAGACACC 2073
 QY 58 -----AsnValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAla 71
 Db 2074 AAGATTTCGCTGCAGACAGACATGCAAATATGACCAACACACC----- 2115
 QY 72 PheLeuLysGluGlnSerThrIleuAlaGlnMetTyrTrpLeuGlnGluIleGlnAsnLeu 91
 Db 2116 -----CTGAACTACGGACCGACCGCCAGGAATTTGATGTGAACCACTTCACAGAACACC 2169
 QY 92 ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAsp 111
 Db 2170 ACTATCAGACGGATATGAAGAGTTGACAGACCTAGAACGGCGACGCTGCCCTGCCAG 2229
 QY 112 LysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLys 131
 Db 2230 GAGCTGAGAGAGTACACACAGATCTGTGGATATGAAACACCACTACAGCGTGGCCACT 2289
 QY 132 ValLysAsnProAspAsnProGlnIleLysLeuLeuGluProGlyLeuAsnGluIle 151
 Db 2290 GGTTCCTCCACCGAATGCG-----AGCTGCTGCGACGCTGACCGCATGTCGACGANTGTG 2343
 QY 152 MetAlaAsnSerLeuAspTyrAsnGlnLysLeuTrpAlaTrpArgLysSerTrpArgSerGlu 171
 Db 2344 ATGGCAGACGCTCCGGAATATGACAGACCTGATGGCATGGGAGGCTGGGAGAACACAC 2403
 QY 172 ValGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGluMetAla 191
 Db 2404 GCGGGAGAGGCGCTCCACAGTTTACCGGAATACGTGGAACCTCATACACAGGCTGCC 2463
 QY 192 ArgAlaAsnHisTyrGlnLysArgTyrGlyLysArgTyrTrpArgLysValLysGlnVal 211
 Db 2464 CGGCTCAATGGCTATGTAGATGACAGGGAGCTGTGGAGTCTATGACAGACACACATCC 2523
 QY 212 ValAspGlyTyrAspTyrSerArgLysLeuIleGluAspValGlnHisThrPheGlu 231
 Db 2524 CTGGAG-----CAAGACCTGAGACGGGCTCTTCCAG 2553
 QY 232 GluIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAla 251
 Db 2554 GAGCTGACGCACTACTCACTCACTCACTGCTAGCTGCTGCGCGGCGCTGACCGCTCAC 2613
 QY 252 Tyr-----ProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuLeuLysPhe 270
 Db 2614 TACGGGGCCCAACATCACTGAGGGGCCCATTCCTGCTCACTGCTGCGGGAACATG 2673
 QY 271 TrpGlyArgPheTrpThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsn 290
 Db 2674 TGGGCGCAGACCTGTCCAAATCATGTAGCTGTGTGCTCCCTTCCCTTCAACCCCTCCG 2733
 QY 291 IleAspValThrAspAlaMetValAspGlnAlaTrpAspAlaArgIlePheLysGlu 310
 Db 2734 ATGGACACCAAGAGGCTATGCTAAAGCAGGCTGAGACCCCAAGAGATGTTTAAAGAG 2793
 QY 311 AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheTrpGluAsn 330
 Db 2794 GCGTATGATTTTCTACGTCCTGGGGGCGTCGCGCGTCCCTCGAGTTCGGAACAG 2853
 QY 331 SerMetLeuThrAspProGlyAsnValGlnLysAlaValCysHisProThrAlaTrpAsp 350
 Db 2854 TCGATCTGAGAGAACGACCAACGCGGAGGTGTGCTGCCAGGCCCTCGGCGCTGGAGC 2913
 QY 351 LeuGlyLysGly-----AspPheArgIleLeuMetCysThrLysValThrMetAspPhe 369

Db 2914 TTCTAACGCGCAAGGACTCCGATCAAGCAATGACGCCACCGCTGACTTGGAGACCTG 2973
 QY 370 LeuThrAlaHisIleGluMetGlyHisIleGlnTyrAspMetAlaTrpAlaIleGlnPro 389
 Db 2974 GTGGTGCCCAACAGAAATGGGCGACATCCAACTATTTCTGACATGACAAAGACTTACT 3033
 QY 390 PheLeuLeuArgAsnGlnAlaAsnGluGlyPheHisGluAlaValGlyLysIleMetSer 409
 Db 3034 GTGGCGCTTGAAGGAGGGTGGCAACCCCGCTTCATAGAGGCCATTTGGGACGTGTAGCC 3093
 QY 410 LeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGln 429
 Db 3094 CTCTCACTGTCTACGGCCCAAGCACTGCACAGCTCACTCACTGCTGAGAGTGGGTGCC 3153
 QY 430 GluAspAsnGlnThrGluIleAsnPheLeuLeuGlnAlaLeuThrIleValGlyThr 449
 Db 3154 AGCGAC---GAGCATGACATCAACTTTCTGATGAATGAGCCCTTACCAAGATGCCCTT 3210
 QY 450 LeuProPheThrTyrMetLeuGluLysTyrTrpArgTrpMetValPheLysGlyLysIlePro 469
 Db 3211 ATCCCTTCAAGTACACTCGTCGATCAAGTGGCGCTGAGGGATTTGATGAGAACATACCC 3270
 QY 470 LysAspGlnTrpMetLysLysTrpTrpGluMetLysArgGluIleValGlyValGlu 489
 Db 3271 AAGGAGACTATTAACCGAGGTGTGGAGCTGAGGCTGAAGTACCAAGGCGCTTCCGCC 3330
 QY 490 ProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSerAsnAsp 509
 Db 3331 CCAAGTCCCAAGACTCAAGTGAATGACTTGAACCCAGGGGCCAAGTTCACATCTCTTACG 3390
 QY 510 TyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnGluAlaLeu 529
 Db 3391 GTGCGTTACATCAGGTACTTGTGACGCTTCATCATCTCAAGTTCCAGTTCCACGAGGACTG 3450
 QY 530 CysGlnAlaAlaLysHisGluGluProLeuHisLysCysAspIleSerAsnSerThrGln 549
 Db 3451 TGGCAGGACAGTGGCCACAGGGGCCCTCACAAGTGTATCTACATCCAGCTCAAGAGAG 3510
 QY 550 AlaGlyLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProThrPheLeuAla 569
 Db 3511 GCGCGGAGCGGCTGGCGACCGCCCATGAGCTGGGCTTCAGTGGCGCGCGGAGAGCC 3570
 QY 570 LeuGluAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGln 589
 Db 3571 ATGAGCTGATCAGGCGCCAGCCCAACATGAGCGCTGGCCATGTGAGACTTCTCAAG 3630
 QY 590 ProLeuPheThrTrpLeuLysAspGlnAsnLys-----AsnSerPheValGlyTrp--- 606
 Db 3631 CCGCTGCTGAGCTGCTCCGACGAGAGACAGCTGATGGGAGAAAGCTGGCTGGCGG 3690
 QY 607 SerThrAspTrpSerProTyrAlaAspGlnSer 617
 Db 3691 CAGTACAACGTGAGCGCAACTCGGCTCGCTCA 3723
 RESULT 20
 AA004027
 ID AA004027 standard; cDNA; 4024 BP.
 AC AA004027;
 XX
 DT 07-SEP-1990 (first entry)
 XX
 DE Human angiotensin converting enzyme (ACE) gene.
 XX
 KW human angiotensin converting enzyme; hypertension;
 KW bradykinin;
 XX
 OS synthetic.
 XX
 XX Key Location/Qualifiers
 FT sig_peptide 23..51
 FT /*tag= a
 FT CDS 23..3944

```

*tag= b
XX XX WO003435-A.
XX PD 05-APR-1990.
XX PF 29-SEP-1989; 89WO-FR00469.
XX PR 27-SEP-1988; 88FR-0012620.
XX (INNM ) INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE.
XX PI Soubrier F, Alhenc-Gelas F, Hubert C, Corvol P;
XX WPI; 1990-058128/08.
XX DR P-PSDB; AAR04111.
XX PR New DNA encoding human angiotensin converting enzyme used eg in
XX diagnosis of hypertension, evaluation of enzyme inhibitors
XX PS Disclosure; P; French.
XX CC This sequence encodes pre-ACE. In the mature protein the 29 amino acid
XX signal peptide is absent. ACE hydrolyses angiotensin I and/or kinins,
XX most notably bradykinin. Vectors contg. the given sequence are used to
XX produce ACE for therapeutic use eg to hydrolyse kinins implicated in
XX inflammation.
SQ Sequence 4024 BP: 857 A; 1263 C; 1175 G; 729 T; 0 other;

Alignment Scores:
Sced. No.: 7.74e-120 Length: 4024
Score: 1337.00 Matches: 255
Percent Similarity: 61.05% Conservative: 118
Best Local Similarity: 41.73% Mismatches: 204
Query Match: 31.16% Gaps: 34
DB: 11 Indels: 9

US-09-978-385-2 (1-805) x AA004027 (1-4024)
OY 20 ThrIleGluGlnAlaIleLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeu 39
DB 1932 ACTGATGAGGCTGAGGAGCCAGCAAGTTGTGAGAGAAATGACCCGACATCCAGGTG 2011
OY 40 PheTyrGlnSerSerLeuAlaSerTyrPasnTyrAsnThrAsnIleThrGlu----- 57
DB 2012 TGGACGAGTATGCCGAGGCCACCTGGAACACTACCAACCAATCCACGAGGACG 2071
OY 58 -----AsnValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerAla 71
DB 2072 AAGATTCTGCTGCAAGAAAGACATGCAATAGCCAAACCCACACC----- 2113
OY 72 PheLeuLysGlnGlnSerThrLeuAlaGlnMetTyrProLeuGlnGlnIleGlnAsnLeu 91
DB 2114 -----CTGAAGTACGCGACCCAGGAGGTTGATGGAACAGTTGCCAAGAACACC 2167
OY 92 ThrValLysLeuGlnLeuAlaLeuGlnLysAsnGlnSerValLeuSerGluAsp 111
DB 2168 ACTATCAGCGGATCATTAAGAGATTGAGACTAGAACGGGAGGCGCTCCGCCAG 2227
OY 112 LysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLys 131
DB 2228 GAGCTGGAGGAGTACAAAGAGATCTGTGGATGTGGAACACCTACAGCGTGGCCACT 2287
OY 132 ValCysAsnProAspAsnProGlnGluCysLeuLeuGlnLysProGlyLeuAsnGlnIle 151
DB 2288 GTGTGCCACCCGATGGC-----AGCTGCTGCAGCTCGAGCCAGATCTGACGAATGTG 2341
OY 152 MetAlaAsnSerLeuAspTyrAsnGluArgLeuTyrPalaTyrGlnSerTyrArgSerGlu 171
DB 2342 ATGGCCACATCCCGGAATATGAGACCTGTTATGGCATGGGAGGCTGGCGAGAACAG 2401
OY 172 ValGlyLysGlnLeuArgProLeuTyrGlnGluTyrValValLeuLysAsnGlnMetAla 191

```

```

DB 2402 GCGGGAGAGCCATCCCACTTTTACCCCAAAATACGTGGAACATCATCAACAGGCTGCC 2461
OY 192 ArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrpArgLysPyrGluValAsnGly 211
DB 2462 CGGCTCAATGSGCTATGTAGATGACAGGGAGACTGTGGAGGCTGTATGTACGAGACCACTCC 2521
OY 212 ValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGlnIleThrPheGlu 231
DB 2522 CTGAG-----CAAGACCTGGAGCGGCTCTTCCAG 2551
OY 232 GlnIleLysProLeuTyrGlnLysLeuHisAlaTyrValArgAlaLysLeuMetAsnAla 251
DB 2552 GAGCTGACAGCCACTTCACTCAACCTGATGCTACGTGGCGGCGCCCTGCACCGCTCAC 2611
OY 252 Tyr---ProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuLeuGlyAspMet 270
DB 2612 TACGGGGCCGACATCACTCAACCTGAGGGGCCATTCTCTGCTCACTGCTGGGAGCATG 2671
OY 271 TrpGlyArgPheTyrPheAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsn 290
DB 2672 TGGCGGAGACCTGTCCAACTATGACTGTGGTGGCCCTTCCCTTCAGCCCTCG 2731
OY 291 IleAspValThrAspAlaMetValAspGlnAlaThrAspAlaGlnArgIlePheLysGlu 310
DB 2732 ATGACACACACAGAGGCTATGCTAAAGCAGGCTGAGCGCCACGAGAGAGATTAAAGAG 2791
OY 311 AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnLysPheTyrGluAsn 330
DB 2792 GCTGATGATTTCTTCACTCCCTCGGGCTGCTGGCCGCTGCTGCTGCTGCTGCAACAG 2851
OY 331 SerMetLeuThrAspProGlyAsnValGlnLysAlaValLysHisProThrAlaTyrAsp 350
DB 2852 TCGATGTGGAGAGACCCAAACAGCGGGGAGGTGCTGCGCACCGCTCGCGCTGGGAGC 2911
OY 351 LeuGlyLysGly---AspPheArgIleLeuMetCysThrLysTyrValThrMetAspPhe 369
DB 2912 TTTCACAAAGCGCAAGGCTTCGCGATCAAGCACTGACCACTGAACTGGAGGACCTG 2971
OY 370 LeuThrAlaHisIleGlnMetGlyHisIleGlnTyrAspMetAlaTyrAlaAlaGlnPro 389
DB 2972 GTGGTGGCCACACAGCAAGATGGGCGACATGCTATTTGCAAGTAAAGACTTACCT 3031
OY 390 PheLeuLeuArgAsnGlnAlaAsnGlnGlyPheHisGluAlaValGlyIleMetSer 409
DB 3032 GTGGCTTGGAGGAGGCTGCCAACCCGCTTCAATGAGGCCATTGGAGCGGTACG 3091
OY 410 LeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGln 429
DB 3092 CTCTCAATGCTACAGCCCAAGACACCTGCACAGTCTCACTGCTGACAGTGAAGGTG 3151
OY 430 GluAspAsnGluThrGlnIleAsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThr 449
DB 3152 AGCGAC---GAGCATGATCAACTTTCGATGATGAATGATGCGCTTGCACAATCGCGCTT 3208
OY 450 LeuProPheThrTyrMetLeuGlnLysTyrPargTrpMetValPheLysGlyGluIlePro 469
DB 3209 ATCCCTTCAGCTACCTCGTGAATGAGTGGCTGAGGCTATTTGATGAGACATCACCC 3268
OY 470 LysAspGlnTrpMetLysTyrTrpGluMetLysArgGluIleValGlyValGlu 489
DB 3269 AAGGAGAACTATTAACAGAGAGGTGGAGCTCAGGCTAAAGTAAAGAGGCTCTGCCCC 3328
OY 490 ProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSerAsnAsp 509
DB 3329 CCAATGCCCAAGCACTCAAGTGAATTTGACCAAGGAGGCAAGTCCATCTCTTACG 3388
OY 510 TyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnIleValLeu 529
DB 3389 GTGCTTACATCGAGTACTTGTGCACTTCAATCATCTCAATTCAGTTCACAGGACAG 3448
OY 530 CysGlnAlaAlaLysHisGlnGluProLeuHisLysCysAspIleSerAsnSerThrGlu 549

```

Db	3449	TGCGAGGACGCTGGCCACACAGGGGCCCCCTGCACAGTGTGACATCTACACAGTCCAAAGAG	3508
Oy	550	AlaGlyGlnLysLeuPheasMetLeuAlaGlyLysSerGluProTrrPThrLeuAla	569
Db	3509	GCCGCGAGGCGCCCTGGGACCGCCCATGAAGCTGGCTTGAAGAGCCCTGGCGCGGAACCC	3568
Oy	570	LeuGlnAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGlu	589
Db	3569	ATGCAGCTGATCACAGGGCCACGCCCATGATGAGCGGCTGGGCATGTGTGACTCTTCAAG	3628
Oy	550	ProLeuPheThrTrpLeuLysAspGlnAsnLys-----AsnSerPheValGlyTyr---	606
Db	3629	CCGCTGCTGCATCTGGCTCCGCCACAGGAGAAAGACCTGCATGGGAGAGAAGCTGGCGTGGCCG	3688
Oy	607	SerThrAspTrpSerProTyrAlaAspGlnSer	617
Db	3689	CAGTACACACTGTGACGCCCACTCCGCTGCTCA	3721

Search completed: March 7, 2003, 07:47:45
Job time : 494 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 7, 2003, 07:29:14 ; Search time 2368 seconds

(without alignments)
5505.645 Million cell updates/sec

Title: US-09-978-385-2

Perfect score: 4291

Sequence: 1 MSSSSWLLSLVAVTAASST.....ISKGNENPGFQNTDVTQTSF 805

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/gen2.1/USPRO.spool/US09978385/runat_28022003_104716_1351/app_query.fasta_1.967
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=20 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09978385.ecgn.1.1.2874 @runat.28022003_104716_1351 -NCP=6 -ICUP=3
-NO_XLPYX -NO_MMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120
-WARM_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlm:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mem:*
24: em_gss_mus:*
25: em_gss_oth:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4249	99.0	3337	11	BC032938
2	1452	93.8	987	9	AL551235
3	1071	25.0	1603	11	AK008530
4	1054.5	24.6	763	13	BI913504
5	1048.5	24.4	879	12	BG401683
6	1031	24.0	766	13	BG622298
7	1022	23.8	676	13	BI61069
8	979	22.0	635	12	BG722079
9	901	21.0	609	13	BI61359
10	864	20.1	565	12	BF789159
11	857	20.0	535	9	AI162058
12	819	19.1	686	10	BB652968
13	781	18.2	452	12	BG428060
14	749.5	17.5	556	13	BM030353
15	730.5	17.0	741	12	BG772384
16	687.5	16.0	800	13	BI26471
17	684	15.9	1154	13	BM460886
18	680	15.8	900	13	BI415293
19	673	15.7	1018	13	BM60537
20	672.5	15.7	778	13	BI831019
21	661.5	15.4	848	13	BI904144
22	659.5	15.4	842	12	BF784147
23	646	15.1	848	13	BI766340
24	643	15.0	492	10	AM260204
25	625	14.6	738	9	AL039738
26	622	14.5	343	13	BI830936
27	615	14.3	877	10	BI196733
28	614	14.3	387	13	AM848579
29	593	13.8	666	13	BI686319
30	588.5	13.7	725	13	BI821905
31	578	13.5	751	13	BI828973
32	575.5	13.4	471	10	AM358360
33	564.5	13.2	823	13	BI914635
34	564	13.1	606	10	AM322773
35	553	12.9	728	10	AM026259
36	549	12.8	459	10	BE138336
37	549	12.8	574	9	AI116021
38	548	12.8	432	10	AM258861
39	547	12.7	865	13	BI825221
40	544	12.7	965	13	BI453005
41	543.5	12.7	775	13	BI103238
42	542	12.6	628	13	BM487251
43	541.5	12.6	724	13	BM290186
44	539	12.6	418	10	BB687602
45	538	12.5	749	10	AM026258

ALIGNMENTS

RESULT 1
LOCUS BC032938 3337 bp mRNA linear HTC 17-JUN-2002
DEFINITION Homo sapiens, angiotensin I converting enzyme (peptidyl dipeptidase A) 2, clone IMAGE:4830668, mRNA.
ACCESSION BC032938
VERSION BC032938.1 GI:21432080
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3337)
AUTHORS Strausberg, R.
TITLE Direct Submission

JOURNAL

Submitted (14-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palakovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: amadansystemsbiology.org
Anup Madan, Jessica Fahney, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whitting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 34 Row: 1 Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 11225608
This clone has the following problem: frame shifted.

FEATURES

SOURCE

1. 3337
/organism="Homo sapiens"
/db_xref="LocusID:59272"
/db_xref="taxon:9606"
/clone="IMAGE:4830668"
/tissue_type="Testis"
/clone_id="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"

BASE COUNT 1023 a 634 c 756 g 924 t
ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3337
Score: 4249.00 Matches: 802
Percent Similarity: 99.75% Conservative: 1
Best Local Similarity: 99.63% Mismatches: 2
Query Match: 99.02% Indels: 1
Gaps: 0

US-09-978-385-2 (1-805) x BC032938 (1-3337)

QY 1 MetSerSerSerSerTripleuLeuSerleuValAlaValThAlaAlaGlnSerThr 20
DB ATGTCAAGCTTCTCCCTGCTCCCTCCACGCTTGTGCTGTAACAGCTGCTCAGTCCACC 99
QY 21 TleGlUGlUGlAlaLysThrPheLeuAspLysPheAsnHisGlAlaGlnAspLeuPhe 40
DB ATTGAGGAACAGCAAGCAATTTGGACAAGTTTAAACGAGCAAGCCGCAAGACCTGTTTC 159
QY 41 TyrGlnSerSerleuAlaSerTriPAsnTyrAsnThrAsnIleThrGlnGluAsnValGln 60
DB TATCAAAAGTTCACCTCTCTTGGAATTTAAACCAATATTACGAAGAAGATGCCAA 219
QY 160 TATCAAAAGTTCACCTCTCTTGGAATTTAAACCAATATTACGAAGAAGATGCCAA 219
QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAla 80
DB AACATGAATTAATGCTGGGCAAAATGCTGCTTTTAAAGAACAGTCCACACTTGCC 279
QY 81 GlnMetTyrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
DB CAAATATATCCACTACAGAAATTCAGAAATTCACAGTACAGTTCAGTTCAGGCTCTT 339
QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
DB CACGCAAAATGGCTCTCAGAGCTCTCAGAAAGCAAGAGCAAGCGTTGAACCAATCTTA 399
QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
DB AATACAAATGAGCACCATCTACAGTACGAAAGTTTGTAAACCAAGATTAATCCACAAAGA 459

QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerleuAspTyrAsnGlu 160
DB TGCCTTATCTTAAACAGAGGTTGAATGAATGAAGCAAGTTTAACTACATGAG 519
QY 161 ArgLeuThrAlaTrpGlnSerTriPAsnSerGluValGlyLysGlnLeuArgProLeuTyr 180
DB AGGCTCTGGGCTGGGAAAGCTGGAGATCTGAGGCTGGCAACAGCTGAGGCAATATAT 579
QY 181 GlnGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAspTyrGly 200
DB GAAGAGATGCTGCTTGAATAATGAGATGGCAAGGCAAAATCATTTATGAGACATGAGG 639
QY 201 AspTyrTriPArgLysAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
DB GATTATTTGAGAGAGAGCTATGAGAAATGAGTATGAGTATGAGTATGAGTATGAGTATG 699
QY 221 GlnLeuIleGlnAspValGlnHisThrPheGlnGlnIleLysProLeuTyrGlnHisLeu 240
DB CAGTTGATGAGATGAGATGAGAAATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 759
QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
DB CATGCTATGTGAGGCGCAAGTTGATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 819
QY 261 CysLeuProAlaHisLeuLeuGlnAspMetTriPArgPheThrPheAsnLeuTyrSer 280
DB TGCCTCTCTCTCATTTGCTTGGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 879
QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
DB TTGACACTTCCCTTGGGCAAGAAACCAACATGATGATCTATCATGATGATGATGATGATG 939
QY 301 AlaThrAspAlaGlnArgGlyIlePheLysGlnAlaGlnLysPhePheValSerValGlyLeu 320
DB GCCTGGGATGACAGAGATTA-TTCAGAGAGGCGGAGAGTCTTTGATCTGTTGGTCTT 998
QY 321 ProAsnMetThrGlnGlnGlyPheTriPAsnSerMetLeuThrAspProGlyAsnValGln 340
DB CCTAATATGACTCAAGATTCCTGGGAAATTCATGCTACAGGACCCAGCAAGATGTTTCA 1058
QY 999 CCTAATATGACTCAAGATTCCTGGGAAATTCATGCTACAGGACCCAGCAAGATGTTTCA 1058
QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlnAspPheArgIleLeuMet 360
DB AAGCAGCTGCTGCTATCCACAGCTTGGGACCTGGGAGGCGGCACTTCAGATCTTATG 1118
QY 1059 AAGCAGCTGCTGCTATCCACAGCTTGGGACCTGGGAGGCGGCACTTCAGATCTTATG 1118
QY 361 CysThrLysValThrMetAspPheLeuThrAlaHisGlnMetGlyHisIleGln 380
DB TGCACAAAGGTGACAAATGAGAGCTTCTGACAGCTATCATGAGATGGGCAATATCCAG 1178
QY 1179 TATGATATGGCATATGCTGCGCAACCTTTCTGCTAGAAATGAGCATATGAGATGATC 1238
QY 401 HisGlnAlaValGlyGlnIleMetSerLeuSerAlaIleThrProLysHisLeuLysSer 420
DB CATGAAGCTGTTGGGGAATCATGCTTTCGACCCACCTTAAGCATTTAAATATCC 1298
QY 421 IleGlyLeuLeuSerProAspPheGlnGlnAspAsnGlnThrGlnIleAsnPheLeu 440
DB ATTGCTCTTCTGTCACCGATTTTCAAGAAAGCAATGAACAGAAATTAATCTCTGCTC 1358
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTrpArg 460
DB AAACAGACACTCAGATGTTGGGACCTGCGCATTTACTTAATCATGTTAGAGAGTGAGG 1418
QY 461 TrpMetValPheLysGlyGlnIleProLysAspGlnTrpMetLysLysTrpArgGlnMet 480
DB TGGATGGCTTTAAAGGGGAAATTCACCAAGACAGAGGAGAAAGTGTGGAGATG 1478
QY 481 LysArgGlnIleValGlyValIleGlnProValProHisAspGlnThrTyrCysAspPro 500
DB AAGCGAGATATGTTGGGCTGTTGGTGAACCTTGCACCATATGATAAATATCTGACACCC 1538

QY 688 ProlysaValSerAspIleIleProArgThrGluValGluTysAlaIleArgMetSer 707
 DB 483 CTTAAATAAGTGTGATATCATTTCTAGAACTGAAAGTGAAGGATCAGATGTC 542
 QY 708 ArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLeuGly 727
 DB 543 CGAGAGCGATCATGATGCTTCCGTCGTGAATGACACAGCTGAGTCTTGGGGATA 602
 QY 728 GlnProThrLeuGlyProProAsnGlnProProValSerIleTyrPheLeuIleValPheGly 747
 DB 603 CACCAACACTGACCTGACCTTACACGCCCCCTGTTTCATATGCGTATGTTTGA 662
 QY 748 ValValMetGlyValIleValValGlyIleValIleLeuIlePheThrGlyIleArgasp 767
 DB 663 GTTGTATGGAGTATGATGAGTGGTGGATGTCATCTGATCTGACTGAGATCAGAGAT 722
 QY 768 ArgTysLysLysLysAsnLysAlaArgSerGlyGluAsnProTyrAlaSerIleAspSer 787
 DB 723 CGGAAAGAAAGAAATAAAGCAAGAGTGGAGAAATCTTATGCTCCATCATATTAC 782
 QY 788 LysGlyGluAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPhe 805
 DB 783 AAAGAGAAATAATATCCAGATTCACAAACATGATGATGTTGACACTCTT 836

RESULT 3
 LOCUS AK008530 1603 bp mRNA linear HTC 19-JAN-2002
 DEFINITION Mus musculus adult male small intestine cDNA, RIKEN full-length
 enriched library, clone:2010305L05;homolog to ANGIOGENSIN
 CONVERTING ENZYME-LIKE PROTEIN (ACE-RELATED CARBOXYPEPTIDASE ACE2),
 full insert sequence.
 AK008530
 VERSION AK008530.1 GI:12842766
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) adult male small intestine cDNA to
 mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library
 clone:2010305L05.
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eumetazoa; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL 99279253
 MEDLINE 10349636
 PUBMED
 REFERENCE
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL 2049374
 MEDLINE 11042159
 PUBMED
 REFERENCE
 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Kashiwagi, K.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.
 Riken integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL 20330913
 MEDLINE 11076861
 PUBMED
 REFERENCE
 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,

Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
 Quackenbush, J., Schriml, L.M., Straub, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barish, G., Blake, J., Boileau, J.D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamlay, M., Lee, N.H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombærts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
 Toyooka, K., Wang, K.H., Weltz, C., Whitaker, C., Wilting, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.
 and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 JOURNAL 21085660
 MEDLINE 11217851
 PUBMED
 REFERENCE
 5 (bases 1 to 1603)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bull, C.,
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
 Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S.,
 Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
 Numazaki, R., Ono, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
 Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Tanaka, T., Tejima, Y., Toyama, T., Yamamura, T., Yamana, I.,
 Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
 Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.go.jp,
 URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for
 further details.

COMMENT
 FEATURES
 source
 location/Qualifiers
 1..1603
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="FANTOM:DB:2010305L05"
 /db_xref="MGI:1902242"
 /db_xref="taxon:10090"
 /clone="2010305L05"
 /sex="male"
 /tissue-type="small intestine"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 49..798
 /note="data source:SPFR, source key:Q9NRA7, evidence:ISS
 homolog to ANGIOGENSIN CONVERTING ENZYME-LIKE PROTEIN
 (ACE-RELATED CARBOXYPEPTIDASE ACB2)
 putative"
 /codon_start=1
 /protein_id="BAB5723.1"
 /db_xref="GI:12842767"
 /db_xref="MGI:1917258"
 /translation="MISGENSEPWTKALENVGARGNDYKPLNYPQPLFDWKQKQNR
 NSFQWNTQSPVADQSIKVISKALGANAAEWENENFELPRSSVAAMKRYFSII
 KNQVPLFEDYRVSCLKPRSPFEPFSPONSDVPRSEVDATIRMSGRINDVFG
 LINDNLEFLGHPLEPYQPPYTHILITSGYVMAALVVGIIILTYTGKGRKKMET
 KRENSYSDMIDIGESNAGFQNSDDQVTS"
 1584..1589
 /note="putative"
 1603
 /note="putative"

polya_signal
 polyA_site

BASE COUNT 465 a 321 c 366 g 450 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 1.44e-118 Length: 1603
Score: 1071.00 Matches: 200
Percent Similarity: 87.508 Conservative: 31
Best Local Similarity: 75.766 Mismatches: 33
Query Match: 24,968 Indels: 0
Gaps: 0

US-09-978-385-2 (1-805) x AK008530 (1-1603)

QY 542 CysAspIleSerAsnSerThrGluAlaGlyIleuPheAsnMetLeuArgLeuGly 561
Db 4 TGAGACATCTCAATTCACCTGAGCTGGCCAGAGCTTCTCAAGATGCTGAGCTTGA 63
QY 562 LysSerGluProThrPheLeuAlaLeuGluAsnValValAlaLysAsnMetAsnVal 581
Db 64 AATTGAGAGCCCTGAGCAAGCCCTGGAAATGTGTGAGAGCAAGCAATATGATGTA 123
QY 582 ArgProLeuAsnThrPheGluProLeuPheThrTrpLeuLysAsnLysAsn 601
Db 124 AATCCACTGCTCATTAATCTTCCACCGCTGTGTGACTGGCTGAAGACAGACGAAAT 183
QY 602 SerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerTrpLysValArg 621
Db 184 TCTTTTGTGGGTGAGCACTGATGAGCCCATATGCCAGCAAGCAATTAAGTGAAG 243
QY 622 IleSerLeuLysSerAlaLeuGluLysPlyAlaTyrGluTrpAsnAspAsnGluMetTyr 641
Db 244 ATTAGCCCTTAATACAGCTCTTGTGACCTAATGATGATGATGATGATGATGATGATG 303
QY 642 LeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGln 661
Db 304 CTGTCCGATCATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 363
QY 662 MetIleLeuPheGlyGluLysPlyAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 681
Db 364 AATGTTCTCTTCTAGAGAGAGATGTAGAGTGTAGAGTGTAGAGTGTAGAGTGTAGAG 423
QY 682 AsnPhePheValThrAlaProLysAsnValSerAspIleTrpProArgThrGluValGlu 701
Db 424 TACTCTTGTGACCTGACCCCAAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 483
QY 702 LysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSer 721
Db 484 GATGCCATCAGATGTCTCTGCGGCGCCATCAATGATGCTTGGCTGGAATGATGATG 543
QY 722 LeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerTrp 741
Db 544 CTGAGATTCTGCGGATTCACCCCAACACTTGAAGCCACCTTCAACGCTCTGTACACATA 603
QY 742 TrpLeuIleValPheGlyValAlaMetGlyValIleValAlaGlyIleValIleLeuIle 761
Db 604 TGGCTGATTAATTTTGGGT 663
QY 762 PheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnProTyr 781
Db 664 GTCACTGGATCAAGGTGCAAGAAAGAAATGAACAAAGAGAGAGAGAGAGAGAGAGAG 723
QY 782 AlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAspVal 801
Db 724 GATCGATGAGCATTTGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861
QY 802 GlnThrSerPhe 805
Db 784 CAGACTTCCTTT 795

RESULT 4
BI913504 763 bp mRNA linear EST 16-OCT-2001
LOCUS BI913504
DEFINITION 6031/9133F1 NIH_MGC_121 Homo sapiens CDNA clone IMAGE:5243048 5',

ACCESSION

BI913504.1 GI:16177911

VERSION

EST.

KEYWORDS

human.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NIH-MGC

REFERENCE

1 (bases 1 to 763)

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: c9abbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LMNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LMNL at:

http://image.llnl.gov

Plate: LAM1613 row: b column: 09

High quality sequence stop: 675.

Location/Qualifiers

1..763

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5243048"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMV-Sport6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH-MGC Library."

BASE COUNT 237 a 167 c 170 g 189 t

ORIGIN

Alignment Scores:

Pred. No.: 4.74e-117 Length: 763

Score: 1054.50 Matches: 225

Percent Similarity: 91.30% Conservative: 6

Best Local Similarity: 88.93% Mismatches: 12

Query Match: 24.57% Indels: 10

DB: 13 Gaps: 3

US-09-978-385-2 (1-805) x BI913504 (1-763)

QY 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaGlnSerThr 20

Db 19 ATGTCAAGCTCTCTGCTGCTCTCTGCTGCTCTCTGCTGCTCTCTGCTGCTCTCTGCT 78

QY 21 Ile-GluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGlnAspLeu 40

Db 79 ATGTGAGGAAAGGCAAGACATTTTGGACAAGTTTAACACGAGAGCGAAGACCTGTT 138

QY 40 eTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGlnAsnValG 60

Db 139 CTATCAAGTTCATCTCTCTGGAATTTTAACACCAATATTAAGTGAAGAAATCTCCA 198

QY 60 nasMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAl 80

Db 199 AAACATGAATAATAGCTGGGACAAATGCTGCTTTTAAGGAAGATCCACACTTGC 258

QY 80 aglMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLysGlnLeuGlnAlaLe 100

Db 259 CCAATATATTCACATCAAGAAATTCAGATATTCACAGCTTCAAGCTGACAGCTCT 318

QY 100 uGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIlele 120
 |||||
 Db 319 TCAGCAAAATGGCTCTTAGGCTCTCGAAGACAGCAAGCGTTGAACAAATCT 378
 QY 120 uAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnG 140
 |||||
 Db 379 AATATCAATGACACACATCTACAGTACGAAAAGTTTGTAAACCAATATATCCACAGA 438
 QY 140 uCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnG 160
 |||||
 Db 439 ATGCTTATTAATCTGAACAGGTTGACTACATTAATGACCAACGTTTACACTACATGA 498
 QY 160 uArgLeuTrrPalatrrp---GluSerTrrPargSerGluValGlyLysGlnLeuArgProle 179
 |||||
 Db 499 GAGGCTCGGCTGGCGCAACCGTGGAGATCTGAGGTGGCAAGCAGCTGAGGCGCAT 558
 QY 179 uTyrGluGluTyrValValLeu-LysAsnGluMetAlaArgAlaAsnHisTyrGluSpt 199
 |||||
 Db 559 AATATGAAGATGTGTGGTGAACAAATGAGATGGCAAGCAAAATCATTATGAGGACT 618
 QY 199 yTgLYAspTrrParg-GlyAspTrrGluVal-AsnGlyVal-AspGlyTrrAspTrrS 218
 |||||
 Db 619 ATGGGATTAATCGAGAGAGAGACATTAAGTACATGCGGTAGCAGTACGCTATGACTAC 678
 QY 218 eTArgGly-GlnLeuIleGlu---AspValGluHisThrPhe---GluGluIleLysPro 235
 |||||
 Db 679 GCCGGGACCGATTGATTCGAACGACTGTGACATACCTTTGACGAGCTTAAACCA 738
 QY 236 LeuTyr-GluHisLeuHisAlaTyr 243
 |||||
 Db 739 TTTACTGAACCTCTTCATGCTTAT 763
 RESULT 5
 BG401683 879 bp mRNA linear EST 12-MAR-2001
 LOCUS 602466350F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4594140 5',
 DEFINITION mRNA sequence.
 ACCESSION BG401683
 VERSION BG401683.1 GI:13295131
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 879)
 AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@bhs-remail.nih.gov
 Tissue Procurement: Clontech Laboratories, Inc.
 cDNA Library Preparation: Clontech Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L10M1335 row: d column: 13
 High quality sequence stop: 619.
 FEATURES
 source
 1..879
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4594140"
 /clone.lib="NIH_MGC_75"
 /lab.host="DH10B (TI phage-resistant)"
 /note="Organ: Kidney; Vector: pDNR-LIB (Clontech); Site_1:
 SfiI (ggcgccctggcc); Site_2: SfiI (ggccattatggcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCGCGAGCGCGACATG-dT(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.65
 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts

by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH MGC Library."

BASE COUNT 275 a 177 c 210 g 216 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 3,15e-116 Length: 879
 Score: 1048.50 Matches: 218
 Percent Similarity: 80.50% Conservative: 9
 Best Local Similarity: 77.30% Mismatches: 12
 Query Match: 24,438 Indels: 45
 DB: 12 Gaps: 4

US-09-978-385-2 (1-805) x BG401683 (1-879)

QY 388 GlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPheHisGluAlaValGlyIle 407
 |||||
 Db 2 CAACCTTTCTGCTAGAAATGAGAACTATGAAAGATTCCATGAGCTGTGGGGAATC 61.
 QY 408 MetSerLeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAsp 427
 |||||
 Db 62 ATGTACCTTTCGACGCCACACCTAAGCATTTAAATCCATGGTCTTGTGACCCGAT 121
 QY 428 PheGlnGluAspAsnGluThrGluIleAsnPheLeuLeuLysGlnAlaLeuThrIleVal 447
 |||||
 Db 122 TTTCAAGAAAGAAATGAAGAAATTAACCTTCCTGCTCAACACAGCCTCAGCATTTG- 180
 QY 448 GlyThrLeuProPheThrTyrMetLeuGluLysTrrPargTrrPheValPheLysGlyLys 467
 |||||
 Db 181 GGGACTCTGCCATTACTTACTACATGTTAGAGAAAGGAGGAGGATGCTTTAAAGGGGA 240
 QY 468 IleProLysAspGlnTrrPheLysLysTrrPheLysLysArgGluIleValGlyVal 487
 |||||
 Db 241 ATTCCCAAGACACCTGATGATAAAAGTGGGAGATGAAG----- 282
 QY 488 ValGluProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSer 507
 |||||
 Db 282 ----- 282
 QY 508 AsnAspTyrSerPheIleArgTyrTrrThrArgThrLeuTyrGlnPheGlnGlu 527
 |||||
 Db 283 -----ACCTTTTACCAATTCAGTTTCAAGAA 309
 QY 528 AlaLeuGlyGlnAlaAlaLysHisGluGlyProLeuHisLysCysAspLysSerAsnSer 547
 |||||
 Db 310 GCACCTTTCAGACAGCTTAACATGAGAGGCCCTCGACAAATGTGCATCTCAAACTCT 369
 QY 548 ThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProTrrPhe 567
 |||||
 Db 370 ACAGAAAGCTGGACAGAAACTGTCAATGTCTGAGGCTTGGAAATCAGAAACCTGGACC 429
 QY 568 LeuAlaLeuGluAsnValValGlyAlaLysAsnMetLysValArgProLeuLeuAsnTyr 587
 |||||
 Db 430 CTGACCTTGGAAAAGTTGTAGAGCAAAAGCAATGATGTAAAGGCGACGCTCAACTAC 489
 QY 588 PheGluProLeuPheThrTrrPheLysAspGlnAsnLysAsnSerPheValGlyTrrPhe 607
 |||||
 Db 490 TTTGAGCCCTTATTTACCTGGCTGGAACACCAAGAAAGATTTCTTTGGGATGGAGT 549
 QY 608 ThrAspTrrPheProTrrAlaAsp-GlnSerIleLysValArgLysSerLeuLysSerAl 627
 |||||
 Db 550 ACCGATGGAGTCCATATGCAAGAACCAAGATCAAGATGAGTAAGCTTAAATTCAGTC 609
 QY 627 AlaLeuGlyAspLysAlaTyrGluTrrPheAspAsnGluMetLeuTrrLeuPheArgSerSerVa 647
 |||||
 Db 610 T---GGAGTTAAGCA-TATGACTGAGC---ACATGAATTTCTCTGTCGATCATCTGT 662
 QY 647 AlaTyr-AlaMetArgGlnTrrPheLysValLysAsnGlnMetIleLeuPheGlyG 667
 |||||
 Db 663 GGCATTTCTTATACGAGTACTTTTAAAGTAAACATCG-----ATGATCTTGGGG 716
 QY 667 Lu 667

```

Db      717 AG 718
RESULT 6
LOCUS   BG962298
DEFINITION 602827090F1 NCI_CGAP_CO24 Mus musculus cDNA clone IMAGE:4981606 5',
          BG962298
ACCESSION BG962298
VERSION   BG962298.1
KEYWORDS  GI:14349935
SOURCE    house mouse.
          Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 766)
AUTHORS  NIH-MGC http://mgi.nci.nih.gov/
TITLE     Unpublished (1999)
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Jeffrey E. Green, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM10983 row: d column: 23
          High quality sequence stop: 708.

FEATURES
     source
         location/Qualifiers
             1..766
             /organism="Mus musculus"
             /strain="FVB/N"
             /db_xref="taxon:10090"
             /clone="IMAGE:4981606"
             /clone_1lb="NCI_CGAP_CO24"
             /lab_host="DH10B ('11 phage-resistant)"/
             /note="Organ: colon; Vector: pCW-SF0R6; Site_1: NotI;
             Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
             Average insert size 1.6 kb. Constructed by Life
             Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 233 a 156 c 179 g 198 t
ORIGIN
Alignment Scores:
Pred. No.: 3,43e-114 Length: 766
Score: 1031.00 Matches: 210
Percent Similarity: 88.85% Conservative: 21
Best Local Similarity: 80.77% Mismatches: 23
Query Match: 24.03% Indels: 9
DB: 13 Gaps: 0

US-09-978-385-2 (1-805) x BG962298 (1-766)
QY 416 LysHisLeuYssSerIleGlyLeuLeuSerProAspPheGlnGlnuAspAsnGluThrGln 435
Db 2 AAGCATCTGAATCCATTGCTCTTCGCCATTCGATTTCACAGATAGCGAAGACAG 61
QY 436 IleAsnPheLeuLeuYsgInAlaLeuThrIleValGlyThrLeuProPheThrTyrMet 455
Db 62 ATAACCTCTCTACTGAAACAGCGCATTTGACATTTGGGAACACTTACCGTTACTTACANG 121
QY 456 LeuGluYsTrpParGTrpMetValPheYsgIleGluIleProLysAspGlnTrpMetLys 475
Db 122 TTACAGAGAGTGAAGTGAATGCTTCGCGGAGCAATTCACCAAGAGAGAGTGAAGA 181
QY 476 LysTrpPheGluMetLysArgGluIleValGlyValAlaGluProValProHisAspGlu 495
Db 182 AAGTGGGAGAGAGGAGCGCGGAGATCGTTGGTGTGGAGAGCTCCCTCATGATGAA 241
QY 496 ThrTyrCysAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyr 515

```

```

Db      242 ACATCTGACCCCTGCATCTCTGTCCATGTTCTAAATGATTCATTCATTCATAT 301
QY 516 TyrThrArgThrLeuYrGlnPheGlnPheGlnGlnAlaLeuYsgInAlaIleHis 535
Db 302 TACACAGAGACCATTTACCAATTCACAGTTTCAAGAGCTCTTTGTCAAGCAGCTAAGTAT 361
QY 536 GluGlyProLeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe 555
Db 362 AATGCTCTCTCACAAATGTGACATCTCAATTCACATGAGCTGGGAGCAAGATGCTCTC 421
QY 556 AsnMetLeuArgLeuGlyLysSerGluProThrPheLeuAlaLeuGlu-AsnValAlaG 575
Db 422 AAGATGCTGAGCTTGGAAATTCAGAGCCCTGGACCAAGCTTGGAAACATGTGGAG 481
QY 575 ValAlaYsAsnMetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPle 595
Db 482 ACCAAGAAATATGATGATGATAAAACCACTGCTCAATTAATCTCAACCGTTGTTGACTGCT 541
QY 595 UlysAspGlnAsnLysAsnSerPhe-ValGly-TrpSerThrAspTrpSerProTyrAla 614
Db 542 GAAGAAGCAGAACAGAAATTTCTTCAGTGGCGGGAACACTGAATGAGACCCCATATAGCC 601
QY 615 AspGlnSerIleLeuValArgIle-SerLeuYsSerAlaLeuGlyAsp-LysAlaTyrG 634
Db 602 GACCAAGCATTTACGATGAGATTAAGCTTAAGATCAGCTCTTGGACCTACATGATATAG 661
QY 634 IuTrpAsnAsp-AsnGluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGln 653
Db 662 ACTGACCAACAGCAAGAAATGCTCTGTCGATCATCTGT-GCATATGCCATGAGAGAC- 719
QY 654 TyrPheLeuYsValLysAsnGluMetIleLeuPheGlyGluLysP 669
Db 720 GTCATGCTCAATATATCAAAAACAGACAGTCTTTTCA-GAGAGAGAT 766

RESULT 7
LOCUS   B1561069
DEFINITION 603254275F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5296531 5',
          B1561069
ACCESSION B1561069
VERSION   B1561069.1
KEYWORDS  GI:15448363
SOURCE    human.
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 676)
AUTHORS  NIH-MGC http://mgi.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
          cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
          Toshiyuki and Piero Carninci (RIKEN)
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM11750 row: f column: 20
          High quality sequence stop: 674.

FEATURES
     source
         location/Qualifiers
             1..676
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone="IMAGE:5296531"
             /clone_1lb="NIH_MGC_97"
             /lab_host="DH10B"
             /note="Organ: testis; Vector: pBluescriptR (modified
             pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag
             ); Oligo-dT primed using primer 5'-TTTCTTTTCTTTTCTT-3',
             size-selected for average insert size 2.2 kb and

```


QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
 DB 459 TGCTTATCTACTGGAACCAAGCTTGAATGAATGCAACAGTTTGAAGCTACCAATGAG 518
 QY 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyGlnLeuArgProLeuTyr 180
 DB 519 AGGCTCTGGGCTGGGAAGCTGAGATCTGAGGTGGCAAGCAAGCTGAGCCATTATAT 578
 QY 181 GluGluTyrValValLeuLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyr 199
 DB 579 GAAGAGTATGTGCTGTTGAAAAATGAGATGCAAGAGCAATCATTTGAGAGACTAT 635
 RESULT 9
 B1561359
 LOCUS 60325121F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5297380 5',
 DEFINITION mRNA sequence.
 ACCESSION B1561359
 VERSION B1561359.1 GI:15448673
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 609)
 AUTHORS NIH-MGC http://mhc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaaps-remail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHM1752 row: j column: 05
 High quality sequence stop: 606.
 Location/Qualifiers
 source
 1. 609
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5297380"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescriptR (modified
 pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTAA-3',
 size-selected for average insert size 2.2 kb and
 normalized to 10^5. This is a primary library enriched
 for full-length clones and constructed using the
 cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC library."
 BASE COUNT 193 a 128 c 137 g 151 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.56e-98 Length: 609
 Score: 901.00 Matches: 186
 Percent Similarity: 97.40% Conservative: 1
 Best Local Similarity: 96.88% Mismatches: 2
 Query Match: 21.00% Indels: 4
 DB: 13 Gaps: 0
 US-09-978-385-2 (1-805) x B1561359 (1-609)
 QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaGlnSerThr 20
 DB 40 ATGTCAAGCTCTTCTGCTGCTTCTGAGCTTATGTCTGTACTCTGCTCAGTCCACC 99

QY 21 IleGluGluAlaValThrPheLeuAspLysPheAsnHisGluAlaGlnAspLeuPhe 40
 DB 100 ATTGAGAACAGAGCCAGACATTTGTGGACAAAGTTTAAACACAGCAAGCCAGCTGTTC 159
 QY 41 TyrGlnSerSerLeuAspTrpAsnTyrAsnThrAsnIlePheGlnGlnValGln 60
 DB 160 TATCAAGTTCACCTGCTTCTTGGAAATTAACACCAATATTACTGAAGAAGATGTCNA 219
 QY 61 AsnMetAsnAsnAlaGlyAspLysTyrPserAlaPheLeu-TyrGlnGlnSerThrLeuAl 80
 DB 220 AACATGATATATCTCTGGGACAAATAGTCTGCTTTTACAAAGAACATCCACTTGC 279
 QY 80 aglMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuAlaLe 100
 DB 280 CCAAAATATACCACTACAGAAATTAATGAAATCTCACAGCAAGCTTCAGCTGAGCTCT 339
 QY 100 uGlnGln-AsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThr-Ile 119
 DB 340 TCAGCAACAAATGGGTCTCTGCTCTCTCAGAAACAAAGAGCAACGGTTGAACACGAAAT 399
 QY 120 LeuAsnThrMetSerThrIleTyrSerThrGlyValCysAsnProAspAsnProGln 139
 DB 400 CTAATATACATATGACCAACATCTACACTGGAAGAGTTTGTACACCAATATCCACA 459
 QY 140 GluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsn 159
 DB 460 GAATGCTTATCTACTGACACAGCTTGAATGAATGAATGAATGAATGAATGAATGAAT 519
 QY 160 GluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyGlnLeuArgProLeu 179
 DB 520 GAAAGCTCTGGGCTTGGGAAAGCTGAGATCTGAGAGTGGGAAG-CAGCTGAGGCCATTA 578
 QY 180 TyrGluGluTyrValValLeuLysAsnGln 189
 DB 579 TATGAGAGATATGTGCTTGAATAATGAG 608
 RESULT 10
 BF789159
 LOCUS 602105037F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4223312
 DEFINITION 5', mRNA sequence.
 ACCESSION BF789159
 VERSION BF789159.1 GI:12094195
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murine; Mus.
 REFERENCE 1 (bases 1 to 965)
 AUTHORS NIH-MGC http://mhc.nci.nih.gov/
 TITLE Unpublished (1999)
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaaps-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHM9811 row: i column: 09
 High quality sequence stop: 676.
 Location/Qualifiers
 source
 1. 965
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4223312"
 /clone_lib="NCI_CGAP_Kid14"
 /lab_host="DH10B (TI phase-resistant)"
 /note="Organ: kidney; Vector: PCMV-SPORT6; Site.1: NotI;

Site-2: Salt: Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI/CCAP Library. 1"

BASE COUNT 286 a 213 c 253 g 212 t 1 others

ALIGNMENT SCORES:

	9.9e-94	Length:	965
Pred. No.:	864.00	Matches:	182
Score:	80.49%	Conservative:	16
Percent Similarity:	73.98%	Mismatches:	45
Best Local Similarity:	20.14%	Indels:	5
Query Match:	12	Gaps:	2

US-09-978-385-2 (1-805) x BF789159 (1-965)

```

QY 1 MetSerSerSerSerTrrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
    |||||
DB 21 ATGTCCAGCTCCTCCTGGCTCTCTCTCAAGCCT-GTTGGCTGTTACTACTGCTCAGTCCCTC 79
    |||||
QY 21 TleGlUGlUGlAlaAlaYthrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
    |||||
DB 80 ACCGGAAGAAATCCAGAACATTTTAAACAACTTAAATCAGGAAGCTGAAGACCTGTCT 139
    |||||
QY 41 TrrGlnSerSerLeuAlaSerTrrpAsnTrrpAsnThrAlaThrGluGluAsnValGln 60
    |||||
DB 140 TATCAAGTTCACCTTCTCTTGGATTTATATATATATATATATATATATATATATATATATAT 199
    |||||
QY 61 AsnMetAsnAsnAlaGlyAspLysTrrpSerAlaPheLeuGluGlnSerThrLeuAla 80
    |||||
DB 200 AAGATGAGTGAAGCTGACCAACCAATGCTCTGCTTTATGAAGAACAGCTTAAGACTGCC 259
    |||||
QY 81 GlnMetTrrpProLeuGlnGluTleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
    |||||
DB 260 CAAAGTCTTCACTACAAAGAAATCCAGACTCCGATCATCAACGCTCACTACAGGCCCTT 319
    |||||
QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrLeuLeu 120
    |||||
DB 320 CAGCAAGAGGCTCTTCACTACTCTCAGCAGACAGAACAAACAGTTGACACATTTCTG 379
    |||||
QY 121 AsnThrMetSerThrTleTrrpSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
    |||||
DB 380 AACACCATGAGACCATTTTACAGTACTGAAAAAGTTTGCAACCCCAAGAACCCCAAGAA 439
    |||||
QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGlnGluMetAlaAsnSerLeuAspTrrpAsnGlu 160
    |||||
DB 440 TGGTTATTACTTGAAGCCAGAGATTGGATGAATTAATGCGCACAGACACTACAACTCT 499
    |||||
QY 161 ArgLeuTrrpAlaTrrpGluSerTrrpArgSerGluValGlyLysGlnLeuArgProLeuTrr 180
    |||||
DB 500 AGGCTCTGGGCAAGGGAGGCTGGAAGGCTGAGTTGGCAACAGCTGAGCGCTGTAT 559
    |||||
QY 181 GluGluTrrpValValLeuLysAsnGluMetAla-ArgAlaAsnHisTrrpGluAspTrrGlu 200
    |||||
DB 560 GAAGAGTATGTGCTCTGAAAAACGAGATGCAAGAACCAATTAATTAACGACTATG 619
    |||||
QY 200 YAspTrrpTrrpArgLysAspTrrpGluValAsnGlyValAspTrrpTrrpArgGlu 220
    |||||
DB 620 GGAATTATTTGGAGAGGGGAGCTATGAACAGAGGAGCAGATGG-TTCACTATTAATACCCGTA 678
    |||||
QY 220 YGlnLeuIleGluAspValGluHisThrPheGlnGluIleLysProLeuTrrpGluHisLe 240
    |||||
DB 679 ACAGTGTATGAGA---TGTGAACGTAATCTCGCAGAAA---TCAAGATTGGTGAGACCTCT 732
    |||||
QY 240 UHisAlaTrrpValArg 245
    |||||
DB 733 TCATGCCCTATGTGAGG 748
    |||||
RESULT 11
LOCUS AA162058 555 bp mRNA linear EST 12-FEB-1997
DEFINITION ms3h11.r1 Stragene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:608613 5' similar to SW:ACE_MOUSE P09470

```

```

ACCESSION AA162058
VERSION AA162058.1 GI:1739055
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 555)
AUTHORS Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Norris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:374045
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 446.
FEATURES
source
1..555
/location/Ovalifera
/oranism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:608613"
/clone_lib="Stragene mouse skin (#937313)"
/sex="Females"
/tissue_type="whole skin"
/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: skin. Vector: pBluescript SK-; Site 1: EORI ; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Whole skin from 11 week old C57BL/6 female mice. Average insert size: 1.0 kb; Uni-ZAP XR vector: -5' adaptor sequence: 5' GATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGCTTTTCTTTTCTTTT 3'"
BASE COUNT 168 a 106 c 150 g 131 t
ORIGIN
Alignment Scores:
Pred. No.: 3.03e-93 Length: 555
Score: 857.00 Matches: 155
Percent Similarity: 89.73% Conservative: 11
Best Local Similarity: 83.78% Mismatches: 19
Query Match: 19.97% Indels: 1
DB: 9 Gaps: 0
US-09-978-385-2 (1-805) x AA162058 (1-555)
QY 126 TleYrrSerThrGlyLysValCysAsnProAspAsnProGlnGluCysLeuLeuGlu 145
    |||||
DB 1 ATTTACACTACTGAAAAAGTTTGCACAAAGAACCAACCAAGAAATCTTATTACTTGAG 60
    |||||
QY 146 ProGluLeuAsnGluIleMetAlaAsnSerLeuAspTrrpAsnGluArgLeuTrrpAlaTrr 165
    |||||
DB 61 CCAGATTGGATGAATTAATGCGCACACACA-GACTTCAACTTGAAGCTCTGGGCATGG 119
    |||||
QY 166 GluSerTrrpArgSerGluValGlyLysGlnLeuArgProLeuTrrpGluTrrpValVal 185
    |||||
DB 120 GAGGCTGAGGCGCTGAGGCTTGGCAAGCAGCTGAGCGCTGTATGAAGATATGTGTC 179
    |||||
QY 186 LeuLysAsnGluMetAlaArgAlaAsnHisTrrpGluAspTrrpGlyAspTrrpArgGly 205
    |||||
DB 180 CTGAAGAAACGAGATGCGCAAGAACCAATTAATTAACGACTATGAGGATTTATGAGAGGG 239
    |||||

```



```

/clone_lib="MARC 2B0V"
/tissue_type="pooled"
/lab_host="DH10B"
/notes="Vector: pCMV SPOR6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, thymus,
semitendinosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
BASE COUNT      178 a      111 c      145 g      122 t
ORIGIN

Alignment Scores:
Pred. No.:      3.59e-80      Length:      556
Score:          749.50      Matches:      142
Percent Similarity: 86.02%      Conservative: 18
Best Local Similarity: 76.34%      Mismatches: 25
Query Match:    17.47%      Indels:      1
DB:             13      Gaps:      1

US-09-978-385-2 (1-805) x BM030353 (1-556)

QY 71 AlAphLeuLysGluInSerThrLeuAlaGlnMetTyrProLeuGlnGluLeuGlnAsn 90
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 2 GCCTTTATGAAAGACAGATCCCGATGCGCAAACTTACTCTGCAAGAAATTCAGAAAT 61
QY 91 LeuThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerValLeuSerGlu 110
    |||||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 62 CTCACACTCAAGGCTCAATTTGAAGGCCCTTCAGCATATGAGCGAGCTCAGCGCTCGGCA 121
QY 111 AsPlySerLysArgLeuAsnThrLleLeuAsnThrMetSerThrLleTyrSerThrGly 130
    :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 122 GAGAGAGCAACAGATTGACACAGATTCTAAATTAATGACACCATCTACAGTACTGGG 181
QY 131 LysValCysAsnProAspAsnProGlnGluCysLeuLeuGluProGlyLeuAsnGlu 150
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 182 AAGATTGAGACCA---AATACAGAGAGAGTCTTACGACTTGAACAGGTTTAATGAC 238
QY 151 IleMetAlaAsnSerLeuAsnPtyrAsnGluArgLeuThrPalatTrpGluSerTrpArgSer 170
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 239 ATTAATGGAAGAAACAGACAGACATCAATACAGAGCTCTGCTGGGAGAGGCTGGAGGCT 298
QY 171 GluValGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGluMet 190
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 299 GAAAGTTGGCAACCACTAAGGCCATTGTATGAGAGTATGAGTCTCTGAAATATGAGATG 358
QY 191 AlaArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrpArgGlyAspTyrGluValAsn 210
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 359 GCAGAGGCCAACAAATTATGAGAGACTACGAGGACTATGAGAGGAGATTATGAGGTGACT 418
QY 211 GlyValAspGlyTyrAspTyrSerArgLysGlnLeuLleGluAspValGluHisThrPhe 230
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 419 GGGGAGGGGAGGACTATGACTACGCGGTGACCAAGTTGATGAAGATGGAACGCACTTT 478
QY 231 GluGluLysProLeuTyrGluHisLeuHisAlaTyrValArgAlaLysLeuMetAsn 250
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 479 GCAGAGATTAAACCATTTATATGACAACTTCATGCTTATGTGAGGGCAAAATTTATGATCAT 538
QY 251 AlaTyrProSerTyrIle 256
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 539 ACCTACCTTCCTATATATC 556

RESULT 15
Bg772384      741 bp      mRNA      linear      EST 15-MAY-2001
LOCUS      602722306F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4839289 5',
DEFINITION      mRNA sequence.
ACCESSION      Bg772384
VERSION      Bg772384.1 GI:14083037
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 741)

```

```

AUTHORS
TITLE      NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL    National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
COMMENT    Contact: Robert Strusberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            Plate: LAM10774 row: k column: 02
            High quality sequence stop: 581.
            Location/Qualifiers
            1..741
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4839289"
            /clone_lib="NIH_MGC_97"
            /lab_host="DH10B"
            /note="Organ: testis; Vector: pBluescript (modified
            pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag
            ); Oligo dT primed using primer 5'-TTTTTTTTTTTTTTTT-3',
            size-selected for average insert size 2.2 kb and
            for full-length clones and constructed using the
            Cap-trapper method (Carninci, in preparation). Library
            constructed by M. Brownstein (NHGRI/NHGR), National
            Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      212 a      188 c      165 g      176 t
ORIGIN

Alignment Scores:
Pred. No.:      1.13e-77      Length:      741
Score:          730.50      Matches:      165
Percent Similarity: 87.88%      Conservative: 5
Best Local Similarity: 85.35%      Mismatches: 11
Query Match:    17.02%      Indels:      13
DB:             12      Gaps:      4

US-09-978-385-2 (1-805) x Bg772384 (1-741)

QY 1 MetSerSerSerSerTyrPheLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 58 ATGTCAAGCTCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 117
QY 21 IleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 118 ATTGAGAACAGGCCAAGACATTTTGGACAAAGTTTAACCGACGACCGAAGACCTGTC 177
QY 41 TyrGlnSerSerLeuAlaSerTyrPasnTyrAsnThrAsnLleThrGluGluAsnValGln 60
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 178 TATCAAGTTCACCTGCTCTCTGGAATTATTAACCAATATTACTGTAAGAGATGCCAA 237
QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluInSerThrLeuAla 80
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 238 AACATGATATATCTGTGGGACAAATGCTGCTCTTAATTAACGAAACAGTCCACACTTGGC 297
QY 81 Gln-MetTyrProLeuGlnGluLleGlnAsnLeuThrValLysLeuGlnAlaLeu 100
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 298 CAATATGATATCACTACCAAGAAATTCAGATATCAAGTCAAGTCAAGTCTCAGGCTCT 357
QY 100 uGlnGlnAsn--GlySerSerValLeuSerGluAspLysSerLys-ArgLeuAsnThrI 119
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 358 TCAGCACACACTGGGTCTTCAGTCTCTCAAGACCAAGACCAACACGTTGAACACAA 417
QY 119 leLeu-AsnThr-MetSerThrLleTyrSerThrGly-LysValCysAsnProAspAsn 138
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 418 TTCTCAATATACACATAGACCATCTACAGTCTGAGACCACTGTATTAACCAAGATATATC 477
QY 138 roGlnGluCysLeuLeuLeuGluProGlyLeuAsnGlu-IleMetAlaAsnSerLeuAsp 157

```

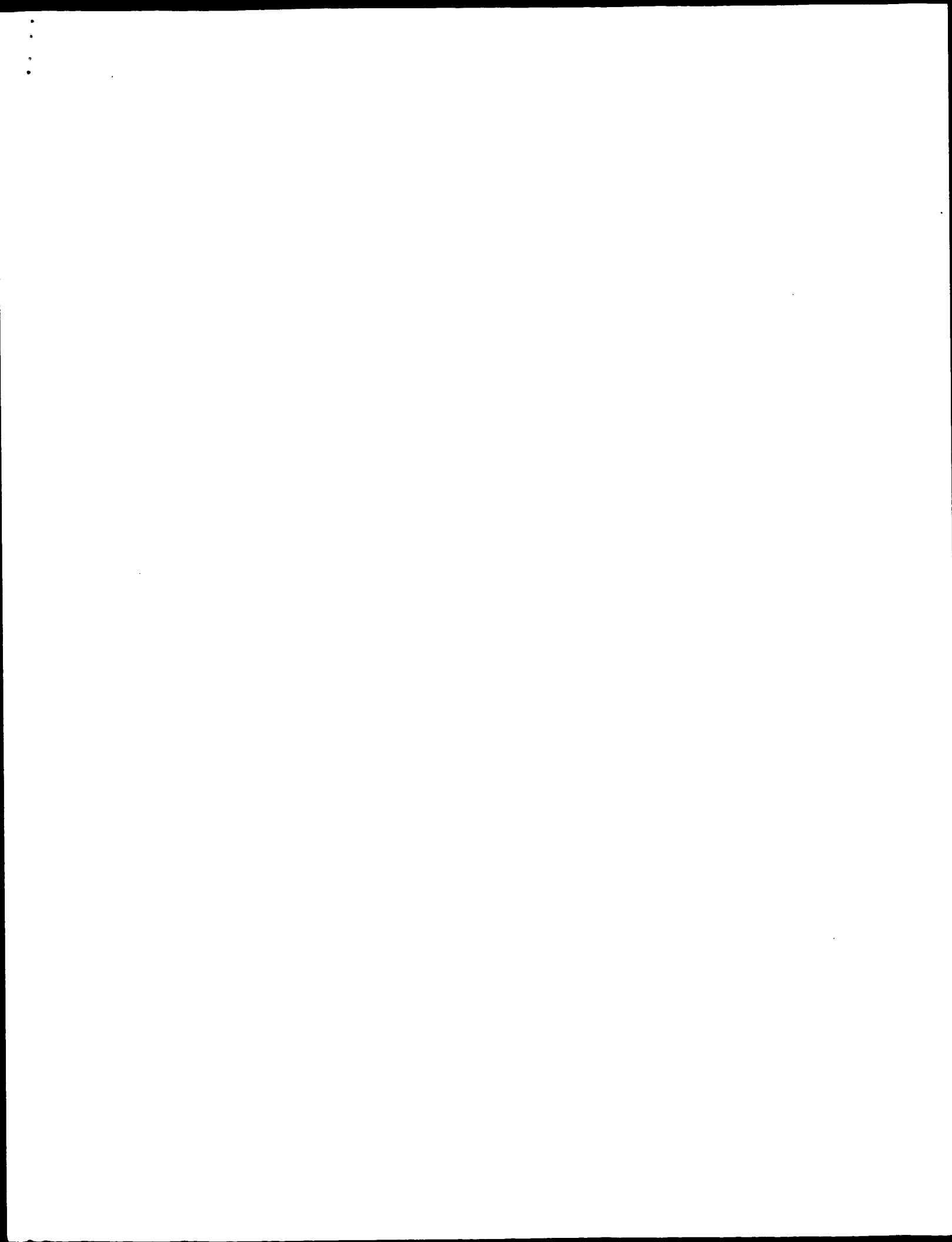
Db 478 CACAGAAATGCTTACTTACCTGACACGCTTGATGACACTGACGCAACAGTTGAC 537
 |||
 Qy 158 TyrAsn---GluArgLeuTrrPalatTrrpJusEr-TrrArgSerGluValGlyLysGlnLe 176
 |||
 Db 538 TACCAACTGACGAGCTCTGGGCTTGGGAAAGCCGTGAGAGACT---GCAGGTCTGTCGACG 594
 |||
 Qy 176 uArg-----ProLeuTyrGlu-GluTyrValValLeu 186
 |||
 Db 595 ACGCTGAGCCACTTATATGAGAGATATGTGGCTCT 632
 |||
 RESULT 16
 B1826471 800 bp mRNA linear EST 04-OCT-2001
 LOCUS B1826471
 DEFINITION 60307617F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5167987 5',
 mRNA sequence.
 ACCESSION B1826471
 VERSION B1826471.1 GI:15938021
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 800)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LAM11417 row: 1 column: 20
 High quality sequence stop: 797.
 Location/Qualifiers
 FEATURES
 source 1..800
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5167987"
 /clone_id="NIH_MGC_119"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site: 1; Note:
 Site: 2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH-MGC Library."
 BASE COUNT 169 a 244 c 230 g 157 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,16e-72 Length: 800
 Score: 687.50 Matches: 128
 Percent Similarity: 68.30% Conservative: 53
 Best Local Similarity: 48.30% Mismatches: 81
 Query Match: 16.02% Indels: 4
 DB: 13 Gaps: 3
 US-09-978-385-2 (1-805) x B1826471 (1-800)
 Qy 230 PheGluGluIleLysProLeuTyrGluHisLeuHisAlaTyrValArgAlaLysLeuMet 249
 |||
 Db 9 TTCACAGAGCTGACGACCTTACCTCAACCTGACCTGCTAGCGCGGCGGCTGCAC 68
 |||
 Qy 250 AsnAlaTyr---ProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuLeuGly 268

Db 69 CGTCACAGAGGGGCCAGCAGCATCAACCTGAGAGGGCCCATTCCTCCTCAGCTGCGGG 128
 |||
 Qy 269 AspMetTrrpGlyArgPheTrrpTrpAsnLeuTyrSerLeuThrValProPheGlyGlnLys 288
 |||
 Db 129 AACATGGGGGCGGAGACCTGCTCCACATCACTAAGACTGATGGTGGCTCCCTTCAGCC 188
 |||
 Qy 289 ProAsnIleAspValTrrAspAlaMetValAspGluAlaTrrAspAlaGlnArllePhe 308
 |||
 Db 189 CCCTCAGTACGACACACAGAGGCTATGCTTAACACAGAGGCTGACGCCAGAGAGATGTT 248
 |||
 Qy 309 LysGluAlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheTrrp 328
 |||
 Db 249 AAGAGGCTGATGATTTCTTCACCTCCGTGGGCTGCTGCCCTCCTGAGTCTTG 308
 |||
 Qy 329 GluAsnSerMetLeuThrAspProGlyAsnValGlnLysAlaValGlyHisProThrAla 348
 |||
 Db 309 AACAGTGCATGCTGGAGAACCCAGCCAGGGGGAGAGGTGGTGGCAGCGCTCGCC 368
 |||
 Qy 349 TrrAspLeuGlyLysGly---AspPheArgIleLeuMetCysThrLysValThrMetAsp 367
 |||
 Db 369 TGGGACTTTCACACAGGCAAGGACTTCGGATCAACAGACAGTGCACACCGTAAGTGGAG 428
 |||
 Qy 368 AspPheLeuThrAlaHisHisGluMetGlyHisIleGlnTyrAspMetAlaTyrAlaAla 387
 |||
 Db 429 GACCTGGTGGTGGCCGCCACAGAAATGGGCCACATCCAGATTTTCATGACAGTCAAAAGAC 488
 |||
 Qy 388 GluProPheLeuLeuArgAsnGlyAlaAsnGluGlyPheHisGluAlaValGlyGluLe 407
 |||
 Db 489 TTTACTGTGGCTTGAAGGAGGGGTGCCAACCCGGCTTCATAGAGCCATGGGAGCTGG 548
 |||
 Qy 408 MetSerLeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAsp 427
 |||
 Db 549 CTACCCCTTCACAGTGTACGCCCAAGACCTCAAGTCAACCTGGTGAACACTGAG 608
 |||
 Qy 428 PheGlnIleLysPheGlnGluThrGluIleAsnPheLeuLeuLysGlnIleAlaThrIleVal 447
 |||
 Db 609 GTGGGACGACAC---GAGCATGACATCACTTCTGTGAAGTGGCCCTTGACAAAGATC 665
 |||
 Qy 448 GlyThrLeuProPheThrTrrMetLeuGluLysTrrPArTrrPmetValPheLysGlyGlu 467
 |||
 Db 666 GCCTTATCCCTTCACACTCACTGCTGATCACTGAGGCGTGGAGGGTATGTAGT-GGAACG 724
 |||
 Qy 468 IleProLysAspGlnTrrPmetLysLysTrrPgluMetLysArgGluIleValGlyVal 487
 |||
 Db 725 ATCCACCAAGAGACTTATACACAGAGTGTGGAGCTCAGCTGAAGAACAGGGCCT 784
 |||
 Qy 488 ValGluProValPro 492
 |||
 Db 785 CTGCCCCCACTGCC 799
 |||
 RESULT 17
 BM460886 1154 bp mRNA linear EST 05-FEB-2002
 LOCUS BM460886
 DEFINITION AGENCOURT_6419622 NCI CGAP OV44 Mus musculus cDNA clone
 IMAGE:5504231 5', mRNA sequence.
 ACCESSION BM460886
 VERSION BM460886.1 GI:18509926
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1154)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabs-remail.nih.gov
 Tissue Procurement: Aaron Hueh
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation

Db 707 CAGCCTCGGCTGGGACTTCTACAAAGGCAAGGACTCCGATCAAGCAGTGCCACCACC 766
 QY 364 ValThMetAspAsp--PheLeuThrAlaHisHisGluMetGlyHisIleGlnTrpAsp 383
 Db 767 GTGAACCTGGAGAGACCTGTTGGTGGCCACACCAATGGCCACATCCACTATTTC 826
 QY 383 etAlaTrpAlaAlaGlnPro-PheLeuLeuArgAsnGlyVal-AsnGluGlyPheHisG 402
 Db 827 TGCAGTACCAAGACTTACCTGTGCTGAGGAGGAGGAGGCCAACCCCGTTCCATGA 886
 QY 402 uAlaValGlyGlu-----IleMetSerLeuSerAlaAlaThrProGlyHisLeuLysSe 420
 Db 887 AGCCCTTGGGGGAGCAGCTGGCTGACCCCTCTCACTGTCTACCCCGCAGCAGCTGCACAG 946
 QY 420 rIleGlyLeu 423
 Db 947 TCTTCAACTG 956
 RESULT 20
 BI831019
 LOCUS BI831019 778 bp mRNA linear EST 04-OCT-2001
 DEFINITION 603080858F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5172423 5',
 mRNA sequence.
 ACCESSION BI831019
 VERSION BI831019.1 GI:15942569
 KEYWORDS Est.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 778)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LLM11429 row: C column: 16
 High quality sequence stop: 770.
 FEATURES
 source
 1..778
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone IMAGE:5172423"
 /clone_id="NIH_MGC_119"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: PCMV-SPORT6; Site_1: NCI;
 Site_2: EcorV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcorV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH_MGC Library."
 BASE COUNT 208 a 159 c 161 g 250 t
 ORIGIN
 Alignment Scores:
 Pred. NO.: 1.38e-70 Length: 778
 Score: 672.50 Matches: 138
 Percent Similarity: 97.24% Conservative: 3
 Best Local Similarity: 95.17% Mismatches: 2
 Query Match: 15.67% Indels: 0
 DB: 13 Gaps: 0

US-09-978-385-2 (1-805) x BI831019 (1-778)
 QY 662 MetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPhe 681
 Db 319 TTGTTCTCTACAGGAGAGAGAGATGTCGAGTGGCTATTGAAACCAAGATCTCCTTT 378
 QY 682 AsnPhePheValThrAlaProLysAsnValSerPheIleIleProArgThrGlyValGlu 701
 Db 379 AATTCTCTTGTCTACTGCACTCAAAATATGTCATATCATATCTAGAACTGAAGTTGAA 438
 QY 702 LysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspSer 721
 Db 439 AAGCCATCAGAGATGCCGAGGCGGTATCATATGCTTCCGTGAATGAACAAC 498
 QY 722 LeuGluPheLeuGlyTleGlnProThrLeuGlyProProAsnGlnProProValSerIle 741
 Db 499 CTAGAGTTTCTGGGATACAGCCAACTTGGACCTCTTACACAGCCCGCTTCCATA 558
 QY 742 TrpLeuIleValPheGlyValValMetGlyValIleValValGlyIleValIleLeuIle 761
 Db 559 TGGCTGATGT--TTTGGAGTTGTATGGAGTGTATGCTTGGCATTTGCATCTCATC 616
 QY 762 PheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnProTyr 781
 Db 617 TTCACCTGGATCAGAGATCGSAGAAAGAAATAAAGCAAGAACTGGAGAAATCTTAT 676
 QY 782 AlaSerIleAspIleSerLysGlyGluAsnAsnProGlyPheGln-AsnThrAspAspVa 801
 Db 677 GCTTCACATCGATATATAGCAAGAAATATATCCAGAGATTCCAAAAACACTGATGATGT 736
 QY 801 IGIThrSerPhe 805
 Db 737 TCACACCTCTCTT 749

Search completed: March 7, 2003, 09:47:36
 Job time : 2390 secs



GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 7, 2003, 07:36:48 ; Search time 91 Seconds
(without alignments)
2712.910 Million cell updates/sec

Title: US-09-978-385-2

Perfect score: 4291
Sequence: 1 MSSSSWLLSLVAVTAQST.....ISKGNPNPQNTDDVQTSF 805

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.model -DEV=xlp
-Q=/cgn2.1/USPFO/US09978385/runat_28022003_104717_1368/app_query.fasta.1.967
-DB=Issued_Patents_NA -OPMT=fastlap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPEXT=0
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=20
-MODE=LOCAL -OUTPMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09978385 -EGN_1_1_61_61 -runat_28022003_104717_1368 -NCPU=6 -ICPU=3
-NO_XLPTX -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEROUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4291	100.0	2415	4	US-08-989-299-3
2	4291	100.0	3396	4	US-08-989-299-1
3	3291	76.7	2350	4	US-09-280-116-40
4	1344	31.3	2478	1	US-08-481-628-1
5	1337	31.2	4020	4	US-09-050-159-130
6	1337	31.2	4024	4	US-09-162-484-18
7	1310	30.5	3942	4	US-09-440-325A-2
8	721	16.8	2082	4	US-08-664-596B-14
9	469	10.9	467	4	US-08-905-223-27
10	381	8.9	848	4	US-09-247-155-27
11	381	8.9	848	4	US-08-157-171-6
12	133.5	3.1	192	1	US-08-157-171-6

13	127	3.0	8700	2	US-08-392-625-16	Sequence 16, Appl
14	127	3.0	8700	2	US-08-466-961A-16	Sequence 16, Appl
15	123	2.9	8700	2	US-08-645-193B-18	Sequence 18, Appl
16	121	2.8	789	4	US-09-280-116-114	Sequence 114, App
17	120.5	2.8	6506	4	US-09-453-702B-1	Sequence 1, Appl
18	120.5	2.8	11384	4	US-08-961-527-45	Sequence 45, Appl
19	120.5	2.8	144	4	US-08-157-171-4	Sequence 4, Appl
20	120	2.8	11303	4	US-08-961-527-115	Sequence 115, App
21	116.5	2.7	2052	3	US-08-630-916A-45	Sequence 45, Appl
22	116.5	2.7	2052	3	US-08-961-527-115	Sequence 45, Appl
23	113.5	2.6	1998	3	US-09-657-481A-3	Sequence 3, Appl
24	113.5	2.6	1998	3	US-08-844-059-1	Sequence 4, Appl
25	113	2.6	1998	4	US-09-431-202-1	Sequence 1, Appl
26	112	2.6	11309	4	US-08-961-527-108	Sequence 108, App
27	112	2.6	1839	3	US-09-120-365-4	Sequence 4, Appl
28	112	2.6	1839	4	US-09-515-039-4	Sequence 4, Appl
29	109	2.5	26385	4	US-08-961-527-202	Sequence 202, App
30	108.5	2.5	9510	4	US-08-961-527-3	Sequence 3, Appl
31	107	2.5	3425	4	US-09-453-702B-256	Sequence 256, App
32	106.5	2.5	3425	4	US-09-107-149-16	Sequence 16, Appl
33	105.5	2.5	37948	4	US-09-251-645-11	Sequence 11, Appl
34	105	2.4	3883	4	US-09-595-424-7	Sequence 7, Appl
35	105	2.4	3883	2	US-08-468-036-33	Sequence 33, Appl
36	105	2.4	3884	4	US-08-376-883-33	Sequence 33, Appl
37	105	2.4	3884	4	US-09-541-782-3	Sequence 3, Appl
38	104	2.4	3884	4	US-09-723-820-3	Sequence 3, Appl
39	104	2.4	3672	1	US-08-491-357-1	Sequence 1, Appl
40	104	2.4	3672	3	US-08-968-633-1	Sequence 1, Appl
41	104	2.4	3672	3	US-09-196-466-1	Sequence 1, Appl
42	101.5	2.4	3672	5	PCT-US96-10823-1	Sequence 1, Appl
43	101.5	2.4	1944	2	US-09-134-001C-316	Sequence 316, Appl
44	101.5	2.4	13977	4	US-08-844-056-1	Sequence 1, Appl
45	101.5	2.4	21706	4	US-09-484-970B-60	Sequence 60, Appl
					US-08-961-527-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1
US-08-989-299-3
Sequence 3, Application US/08989299
Patent No. 6194556
GENERAL INFORMATION:
APPLICANT: Robinson, Susan L.
TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: POLEY, HOAG & ELLIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,299
FILING DATE: 11-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Arnold E., Beth
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MIA-025.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-7000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2415 base pairs

Db 1921 TACCTGTCGATCATCTGTTCATATGCTATGAGGACGACTTCTTTAAAGTAAAAAT 1980
 QY 661 GlnMetIleuPheGlyGluAspValArgValAlaAsnLeuLysProArgIleSer 680
 Db 1981 CAGATGATCTTTTGGGAGGAGATGCGAGTGGCTAAATTTGAAACCAAGAAATCTCC 2040
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700
 Db 2041 TTTAATTTCTTGTCTACGTCACCTTAAAAATGTCTCTGATATATCTCTTGAACCTGAAGTT 2100
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
 Db 2101 GAAAAGCCATCAGAGATGTCGCCGAGCCCTATCAATGATGCTTCGCTGATGACCAAC 2160
 QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
 Db 2161 AGCTTAGAGTTCTGGGGATACAGCCAAACACTTGACCTCCCAACAGCCCTCTTTC 2220
 QY 741 IleThrPheIleValPheGlyValValMetGlyValIleValIleValIleValIleVal 760
 Db 2221 ATATGGCTGATGTTTGTGAGTTGTGATGGAGATGATGCTGTGCTTCATCTCG 2280
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780
 Db 2281 ATCTTCACCTGGATCAGATCGAAGAAAGAAATAAAGCAAGAGTGAAGAAATCTCT 2340
 QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800
 Db 2341 TATGGCTCATGATATTGCAAGAGAAATAATTCAGATTCACAAACACTGATAT 2400
 QY 801 ValGlnThrSerPhe 805
 Db 2401 GTTCAGACCTCTT 2415

RESULT 2

; Sequence 1, Application US/08989299
 ; Patent No. 6194556
 ; GENERAL INFORMATION:
 ; APPLICANT: Action, Susan L.
 ; TITLE OF INVENTION: ANGIOGENESIS CONVERTING ENZYME HOMOLOGY
 ; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY, HONG & ELIOT LLP
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109-2170
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/989,299
 ; FILING DATE: 11-DEC-1997
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Arnold E., Beth
 ; REGISTRATION NUMBER: 35,430
 ; REFERENCE/DOCKET NUMBER: MTA-025.01
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-832-1000
 ; TELEFAX: 617-832-7000
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3396 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 82..2496
 US-08-989-299-1

Alignment Scores:
 Pred. NO.: 0
 Score: 4291.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 4
 Gaps: 0

US-09-978-385-2 (1-805) x US-08-989-299-1 (1-3396)

QY 1 MetSerSerSerThrPheLeuLeuSerLeuValAlaThrAlaIleGlnSerThr 20
 Db 82 ATGTCAACCTCTTCTGCTCTCTCTCTCAAGCTTGTCTGTACTGCTGCTCAGTCCACC 141
 QY 21 IleGluGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
 Db 142 ATTAGGCAACAGGCCAAGACATTTTGGACAAAGTTTAACCAAGAGCCGAAGACCTGTTTC 201
 QY 41 TyrGlnSerSerLeuAlaSerThrPAsnTyrAsnThrAsnIleThrGluAsnValGln 60
 Db 202 TATCAAGTTACTTCTGCTCTTGGAAATTAACCAATATATCTGTAAGAGAAATGTCCAA 261
 QY 61 AsnMetAsnAsnAlaGlyAspLysThrSerAlaPheLeuLysGluGlnSerThrLeuAla 80
 Db 262 AACATGAATATATGCTGGGCAAAATGCTGCTCTTTTAAAGAGACGTCACACATTTGCC 321
 QY 81 GlnMetThrProLeuGlnIleGlnAsnLeuThrValLysLeuGlnIleAlaLeu 100
 Db 322 CAATGTATCCCTACCAAGAAATTCAGAAATCTCACAGCTTCAAGCTGACGCTCTT 381
 QY 101 GlnGlnAsnGlySerSerValIleuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
 Db 382 CAGCAAAATGGCTTTTCAAGTCTGCTGTCAGAAAGACAGCAAGAGGTGAACACAAATTCGA 441
 QY 121 AsnThrMetSerThrIleLysSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
 Db 442 AATACAAATGAGCAGACACTACACTGAGTGAAGAAAGTTGTAAACCAATATCCACAAGAA 501
 QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspLysGlu 160
 Db 502 TGCTTATTACTTGAAACGAGTTTGAATGAATATGCAACAGTTTACCTCAATATGAG 561
 QY 161 ArgLeuThrAlaTrpGlnSerThrPArgSerGluValGlyLysGlnLeuArgProLeuTyr 180
 Db 562 AGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGGTGGCAAGCAAGCTGAGGCCATTATAT 621
 QY 181 GluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspLysGly 200
 Db 622 GAGAGATATGCTGCTTGAAGAAATGAGATGCAAGAGCAAAATCATTAATGAGGACTATGG 681
 QY 201 AspTyrThrPArgLysPyrGluValAsnGlyValAspGlyLysTyrAspLysSerArgGly 220
 Db 682 GATTATTGGAGAGAGACTATGAAGTAAATGGGGTGAATGCTATGACTACACCCGCGGC 741
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240
 Db 742 CAGTGTATGAAGATGGAGAACATACCTTTGAAGAGATTAACCATTAATATGAACATCTT 801
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
 Db 802 CATGCCATATGAGGCAAAAGTGAATGCAATGCTATCTCCATATVCAAGTCAATATGA 861
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetThrGlyArgPheThrPheHisLeuTyrSer 280
 Db 862 TGCCCTCCCTCATTTGCTGTGATATGTGGGGTAGATTTTGGCAAAATCTGTACTCT 921
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300

```

Db 922 TTGACAGTTCCTTTGGACAGAAACCAACATAGATGTTACTGATGCAATGGTGGACAG 981
Qy 301 AATATPASPAlaGlnArgIlePheLysGluAlaGluLysPheValSerValGlyLeu 320
Db 982 GCGTGGATGACAGAGATATTTCAAGAGGCGGAGAAAGTTCTTGTATCTGTGTGCTT 1041
Qy 321 ProAsMetThrGlnGlyPheThrProLysAsnSerMetLeuThrAspProGlyLysValGln 340
Db 1042 CCTAAATATGACTCAAGGATTCGGGAAATTCATGCTAAACGACCCAGAAATATTTGAG 1101
Qy 341 LysAlaValAlaCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1102 AAAGCAGTGTGCATGCCATCCACACTGGGAGCTGGGGAAGGGGACTTCAGATTCCTTATG 1161
Qy 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380
Db 1162 TGCACAAAGGTGACATGAGACACTTCCTGACAGCTCATGAGATGGGCAATATCCAG 1221
Qy 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
Db 1222 TATGATATGTCATATGCTCACAACCTTTGCTTGAAGAAATGAGCTATATAGAGATTC 1281
Qy 401 HisGluAlaValAlaGlyLysIleLeuMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
Db 1282 CAGGAGCTGTGGGAAATCATGTCACCTTCTGCAGCCACACCTAAGCATTTAAATCC 1341
Qy 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluLysAsnPheLeuLeu 440
Db 1342 ATTTGCTCTTCTGTCCACCATTTTCAAGAGCAATGAAAGAAATAACTCTGTGCTC 1401
Qy 441 LysGluAlaLeuThrIleValAlaGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460
Db 1402 AAACAAGCACTCCAGATGTTGTGGAGCTTGCCATTTACTTACATGTTAAGAGAGGAG 1461
Qy 461 TrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTrpPgluMet 480
Db 1462 TGGATGCTCTTTAAAGGGGAAATCCCAAGAGCACTGATGATAAAAGGGTGGGAGATG 1521
Qy 481 LysArgGluIleValAlaGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500
Db 1522 AACCGAGAGATAGTGGGTGGTGGAGACTGTGCCCATGATGAAACATACAGTGTGACCCC 1581
Qy 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520
Db 1582 GCATCTCTGTCCATGTTCTATATGATTTACTCATTTGCATATGATACCAAGGACCCCT 1641
Qy 521 TyrGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaAlaLysHisGluGlyProLeuHis 540
Db 1642 TACCAATTCACAGTTTCAAGAGCACTTTGTCAAGCAGCAATGAAACATGAAAGCCCTG 1701
Qy 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560
Db 1702 AAATGTGACATCTCAAACTCTACAGAGAGCTGGACAGAAACTGTCTCAATTTGCTGAG 1761
Qy 561 GlyLysSerGluProTyrThrLeuAlaLeuGluLysAsnValValGlyAlaLysAsnMetAsn 580
Db 1762 GGAATAATCAGAACCTCGAGCCCTAGCATTTGGAATAATGTTGTAGAGCAAAAGAACAT 1821
Qy 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAsnGlnLys 600
Db 1822 GTAAAGGCACTGCTCAACTTCTTGAAGCCCTTATTTACCTGGGTGAAAGCCGAGAACAG 1881
Qy 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620
Db 1882 AATTTCTTTTGTGGATGGATGACGACTGAGTCCATATGACAGACCAAGACATCAAAAGTG 1941
Qy 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet 640
Db 1942 AGGATTAAGCCTTAATCAGCTTTGGAGATTAAGCATATGAGCAACCATGAAATG 2001
Qy 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660

```

```

Db 2002 TACCTGTTCCGATCATCTGTGCATATGCTATGAGCAGACTTTTAAAGTAAATAAT 2061
Qy 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680
Db 2062 CAGATGATCTCTTTTGGGAGGAGATGTCGAGTGGCTAATTTGAAACCAAGAACTCTCC 2121
Qy 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleLeuProArgThrGluVal 700
Db 2122 TTTAATTTCTTTGCTGACTGCACCTAAATATGCTGCTGATATCATTTCTTCAACTGAAG 2181
Qy 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2182 GAAAGGCCATCAGAGATGTCGCCGAGCGTATCATATGATCTTCCCTGTAATGACAAAC 2241
Qy 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2242 AGCCTAGGTTTCTGGGAGATACACCAACACTTGGACCTCTTAAACCGCCCCCTGTTTC 2301
Qy 741 IleThrPheIleValPheGlyValAlaMetGlyValIleValAlaGlyIleValIleLeu 760
Db 2302 ATATGGCTGATGTTGTTTGGAGTTGTGATGGAGATGATAGTGGTGGCATTTGATCCTG 2361
Qy 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnPro 780
Db 2362 ATCTTCACTGGGATCAGAGATCGGAGAGAGAAATTAACCAAGAGCTGAGAAATCCCT 2421
Qy 781 TyrAlaSerIleAspIleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAsp 800
Db 2422 TATGCTCTCATCGATATTTAGCAAGAGAGAAATATATCCAGATTTCCAAACACTGATG 2481
Qy 801 ValGlnThrSerPhe 805
Db 2482 GTTCAGACCTCTCTT 2496

RESULT 3
US-09-280-116-40/c
Sequence 40, Application US/09280116A
Patent No. 6331427
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OR INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
FILE REFERENCE: 5800-24, 035800/176965
CURRENT APPLICATION NUMBER: US/09/280,116A
NUMBER OF SEQ ID NOS: 268
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 40
LENGTH: 2350
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
OTHER INFORMATION: angiotensin-converting enzyme
US-09-280-116-40

Alignment Scores:
Pred. No.: 0 Length: 2350
Score: 3291.00 Matches: 726
Percent Similarity: 92.62% Conservative: 2
Best Local Similarity: 92.37% Mismatches: 12
Query Match: 76.70% Indels: 46
DB: 4 Gaps: 2

US-09-978-385-2 (1-805) x US-09-280-116-40 (1-2350)
Qy 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaVal-ThrAlaAlaGlnSerTh 20
Db 2305 ATGTAAACCTCTCTCTGCTCTCTCTCTGCTGCTGTGCTGTGCTGCTGCTGCTGCTGCTG 2246
Qy 20 rIleGluGluGlnAlaLysThr-PheLeuAspLysPheAsnHisGlu-AlaGluAspLeu 39
Db 2245 CATTGAGAGACAGCCAGACATTTTGGACAGATTAAACAGGAAAGCCGAGAGACTG 2186
Qy 40 PheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThr-AsnIleThrGluLysAsn 59

```

Db	2185	TTCTATCAAGTCACTGCTCTTGGAAATATAACCCCAATTTACTGAAAGATCT	2128
Qy	59	IGLHSMNETHLSMNSALAGLYASP-LysTrpSer-AlaPheLeuLysGluGlnSerThr	78
Db	2125	CCAAAGACGAAATATATGCTGGGACAAATATGGTGTGCTTTTAAAGACAGTCCACA	2066
Qy	79	LeuAlaGlnMetCysTrpProLeuGlnGluLeuGlnSerLeuThrValLysLeuGlnLeuGln	98
Db	2065	CTTGGCCCAATATGATATCCACTACACAGAAATATCAAGATATCCACACTCAAGCTTATCTGACG	2006
Qy	99	AlaLeuGlnGlnGlnSerLysSerValLeuSerGluAspLysSerLysTrpLeuLysThr	118
Db	2005	GCTCTTCAACAAATATGGTCTTCACTGCTCTCAGAACACAAAGCAACGTTGTAACACA	1946
Qy	119	IleLeuAsnThrMetSerThrIle-TyrSerThrGly-LysValCysAsnProAspAsn	138
Db	1945	ATTCTAATATACATATGACGCCACCACTACAGATGACGAAAAAGTTGTATACCAAGATATC	1886
Qy	138	roGlnGluCysLeuLeuLeuGlnProGlyLeuAsnGluIleMetIleAsnSerLeuAsp	158
Db	1885	CACAGAAATCTTATTAATCTTACCTGCAACAGAGTTGAATGAATATATGCAACAGTTTAAC	1826
Qy	158	lysAsnGluArgLeuLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgp	178
Db	1825	ACATGAGAGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGTCCGGCAAGCAGCTGAGGC	1766
Qy	178	roleuTrpGluGluTrpValValLeuLysAsnGluMetIleArgIleAsnHisTrpGlu	198
Db	1765	CATTATATATAAGATATGCTGTGTAATAATAGATGCGCAAGACCAATATATTATGAG	1706
Qy	198	spTrpLysAspTrpTrpArgLysAspTrpGlu-ValAsnGlyVal-AspGlyLysAsp	217
Db	1705	ACATGGGATTTATGGAGAGAGACATATGAAAGTAATAGGGGTAAAGTGCATATGCTA	1646
Qy	217	r-SerArgLysGlnLeuIleGluAspValGluHisThr-PheGluGluIleLysProLeu	236
Db	1645	CATGCCCGCGCCAGTGTATGAAGATGTGGAAACATACCTTTTGAGAGATTTAAACCATTA	1586
Qy	237	TYRGLHLSLSLEHLSALATYRVALAVALALysLeuMetAsnAlaTrpProSerTrpIle	256
Db	1585	TATGACATCTTCAATGCTTATGTGAGCGCAAAATGATGATGCTTCTCTTCAATAC	1526
Qy	257	SerProIleGlyCysLeuProAlaHisLeu-LeuGlyAspMetTrpGly-ArgPheTrp-	275
Db	1525	ACTCCAAATTTGATGCCCTCCCTGCTCATTTCCGTTGGTGAATATGGGGTTAAGATTTGGG	1466
Qy	276	ThrAsnLeuTrpSerLeuThrVal-ProPheGlyGln-LysProAsnIle-AspAlaThr	294
Db	1465	ACAAATCTGATCTTTGACAGATGCTCCCTTGGACAGAGAAACCAACATACATATTACT	1406
Qy	295	AspAla-MetValAspGlnAlaTrpAspAla-GlnArgIlePheLysGluAla-GluLys	313
Db	1405	GATGCCAATGTGTGACAGAGTGCACCAAGAAATTTCAAGAGGCGCCAGAAAG	1346
Qy	314	PhePheValSerValGlyLeuProAsnMetHisGlnGlyPheTrp-GluAsnSerMetIle	333
Db	1345	TTCTTTTATCTGTGGTCTCTTAATATGACTCAAGATTTCTGGGAAATTTCAATGCT	1286
Qy	333	uThr-AspProGlyAsnValGlnLysAlaValCysHisProThrAlaTrpAspLeu-Gly	352
Db	1285	AAGCGACCCAGAAATGTTCAAGAAAGAGCTGCGCATCCACACTTGGGAGCTGGGGG	1226
Qy	353	LysGlyLysP-PheArgIle-LeuMetCysThrLysVal-ThrMetAspAspPheLeuThr	372
Db	1225	AAGGGCGACTTCAGAGATCTTATATGTCACAAAGGTGACATGACACACTTCTGTACAG	1166
Qy	372	IaHisHisGluMetGlyHisHisIleGlyLysPheMetAlaTrpAlaIaGlnProPheLeu	392
Db	1165	CTCATCATATAGATGGGCGCATATCCAGATATATATGGCATATGCTGCACAACTTTTCTGG	1106
Qy	392	euaRgsnGlyAlaAsnGluGlyPheHisGlnIleValGlyIleIleMetSerLeuSerA	412

Db	1105	TAAAGAAATGGAGCTTAATGAAGATTCCATGAAGCTGTTGGGGAATCATGTCACTTTCG	1046
Qy	412	laalathPrC-LysHisLeuLysSerIleGlyLeuLeu---SerProAspPheGlnGlu	430
Db	1045	CAGCCACACCTTAAGACATTAAATATCCATTCGTTGTTTTTTTGTCCACCCGAAATTTTTCAGAA	986
Qy	431	AspAsnGluThrGluIleAsnPheLeuLeuLysGln-AlaLeuThrIleValGlyThr-L	450
Db	985	GACATTCGAACAGAAATAAATCTCCGCTCAACAAAGACATCAGCATGTGGGACATCC	926
Qy	450	eunProPheThrTYrMetLeuGluLysTPRATGTPRMeValPheLysGlyGluIlePOL	470
Db	925	TGCCATTACTTACATGTTTAGAAGATGGAGTGATGGCTTTTAAAGGGGAAATTTCCA	866
Qy	470	ysaSPGlnTPRmetLys-LysTPRtpRgluMetLysArgGluIleValGlyValI---	488
Db	865	AAGACACAGTGAATGAACAAAGTGCGGAGATGAAGCAGAGATAGTTGGGTGGTTGGG	806
Qy	489	GluProVal-ProHisAspGluThrTYrCysAspProAlaSerLeuPheHisValSerAs	508
Db	805	AACTTGTGCCCATGATGAAGAAACATCTGTGACCCCGCATCTCTGTTCCATGTTTCTAA	746
Qy	508	naSPtyrSer-PheIleArgTYrTYrThrArgThrLeuTYr-GlnPhe-GlnPheGlnI	527
Db	745	TGATTACTCCATTCATTCATTCATTCACAAAGAACCTTTTAAACCATTCACCATTCAGA	866
Qy	527	uAlaLeuCys-GlnAlaAlaLysHisGluGlyProLeuHisLysCysAspIleSer-Asn	546
Db	685	AGACCTTTTGTCACACCAGCTAAACATGAAGCCCTTGCAAAATGTCATCTCGTAAC	626
Qy	547	SerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProtp	566
Db	625	TCTACAGAAAGCTGGACAGAACTGTCATATGCTGAGCTTGGAATTCAGAACCTGG	566
Qy	567	ThrLeuAlaLeuGluAsnValValGlyAlaLys-AsnMetAsnValArgProLeuLeuAs	586
Db	565	ACCTTGACATTTGGAAATGTTGTAGGAGCAAAAGAACATGATGAAGGCCACATGCTCAA	506
Qy	586	n-TyrPheGln-ProLeuPheThrTPRLeuLysAspGlnAsnLysAsnSerPheValGly	605
Db	505	CGTACTGTGAGCGCGTTATTACTGCGTGAAGACAGCAAGAAATCTTTTGTGGGA	446
Qy	606	TrpSerThrAspTPRSerProTYrAlaAspGln-SerIleLysValArgIle-SerLeuL	625
Db	445	TGGAGTACCGAGCTGAGATTCATATGCAGACCAAAAGCATCAAACTGAGCATTAAGCTTA	386
Qy	625	ysSerAlaLeuGluIAspLysAlaTYrGluTPRAsnAspAsnGluMetTYrLeuPheArgS	645
Db	385	AATTCAGCTCTTGAGGATTAAGCATATGAATGAAGACAGCAAAATGATACCTGTCGAT	326
Qy	645	eTSerValAlaTYrAlaMetArgGlnTYrPheLeuLysValLysAsnGln-MetIleLeu	664
Db	325	CATCTGTGTCATATGCTATGAGGACAGTACTTTTAAAGTAATAAATCAGCATGTTCTT	266
Qy	665	PheGlyGluGlu-AspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePh	684
Db	265	TTTGGGAGAGAAAGATGTGCGAGTGGCTAAATTTGAAACCAAGANTCTCCTTAATTTCTT	206
Qy	684	e-ValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGlyLysAlaI	704
Db	205	TGTGCTACGTGCCCTTAATAATATGCTGATATCATCTCTGAAGTGAAGTTGAAGGCA	146
Qy	704	LeaYmetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluP	724
Db	145	TCAAGATGTCCGGAGACGATCATATATGCTTCCGTGATATCAACAACAGCCTTAGAGT	86
Qy	724	heLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer	740
Db	85	TTCTGGGATACAGCCAAACACTGGAGCTCTTAACCAAGCCCTCTGTTCC	36
RESULT 4			
US-08-481-626-1			
Sequence 1, Application US/08481626			

Patent No. 5801040
 GENERAL INFORMATION:
 APPLICANT: Soubrier, Florent
 APPLICANT: Albert-Gelas, Francois
 APPLICANT: Hubert, Christine
 APPLICANT: Corvol, Pierre
 TITLE OF INVENTION: Nucleic Acid Coding for the Human
 TITLE OF INVENTION: Testicular Angiotensin Converting Enzyme (ACE) and its
 TITLE OF INVENTION: Uses, Especially for the In Vitro Screening for this
 TITLE OF INVENTION: Enzyme in the Organism
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Flanagan, Henderson, Farabow, Garrett &
 ADDRESS: 1300 I Street, N.W., Suite 700
 CITY: Washington
 STATE: D.C.
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US-08/481,626
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/656,183
 FILING DATE: 04-MAR-1991
 CLASSIFICATION: 435
 APPLICATION DATA:
 APPLICATION NUMBER: FR 89-09062
 FILING DATE: 05-JUL-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 REFERENCE/DOCKET NUMBER: 04958-0006-02000
 TELEPHONE: 202-408-4000
 TELEFAX: 202-408-4400
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2478 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-481-626-1
 Alignment Scores:
 Pred. No.: 2,29e-148 Length: 2478
 Score: 1344.00 Matches: 259
 Percent Similarity: 60.97% Conservative: 119
 Best Local Similarity: 41.77% Mismatches: 204
 Query Match: 31.32% Indels: 38
 DB: 1 Gaps: 10
 US-09-978-385-2 (1-805) x US-08-481-626-1 (1-2478)
 QY 15 ThrAlaAlaGlnSer-----ThrlleGluGluGlnAlaAlaLysThrPheLeuasp 30
 Db 209 ACATCGCCCAAGCAACCTGCTGATGATGAGGCGCCACCAAGTTGTGGAG 268
 QY 31 LysPheAsnHisGluAlaGlnAspLeuPheTyrGlnSerSerLeuAlaSerTrrPasnTyr 50
 Db 269 GAATATGACCGACATCCAGGTGTGTGAACGATGTGCCGAGGCCAATGGAGCTAC 328
 QY 51 AsnThrAsnIleThrGluGlu-----AsnValGlnAsnMet 62
 Db 329 AACACCAACATCCACACAGACAGCAAGATTCTGTGACAGAGAACATGCAATAGCC 388
 QY 63 AsnAsnAlaGlyAspLysTrrPserAlaPheLeuLysGluGlnSerThrLeuAlaGlnMet 82

Db 389 AACCAACAC-----CTGAGTACGAGCCAGCAGAG 424
 QY 83 TyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeuGln 102
 Db 425 TTGTATGTGAACAGTGTGCAACACCACTATCAACCGGATATGAAGAGGTTCAGAGC 484
 QY 103 AsnGlySerValLeuSerGluAspLysSerTyrArgLeuAsnThrIleLeuAsnThr 122
 Db 485 CTAGAACGGGCGAGGCTGCTGCCAGAGAGCTGAGAGTACACAGATCTGTGGAT 544
 QY 123 MetSerThrIleTyrSerThrGlnGlyValCysAsnProAspAsnProGlnLeuLysLeu 142
 Db 545 ATGGAACCACTACAGAGCTGCGCCATGTGCTGCACCGAATGCG-----AGCTGCTG 598
 QY 143 LeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspLysArgLeu 162
 Db 599 CAGCTCGAGCCAGATCTGACGAGATGTATGGCCATCCCGGAATATGAGACGCTTGA 658
 QY 163 TrpAlaTrrPgluSerTrrPargSerGluValGlyGlnLeuArgProLeuTyrGluGlu 182
 Db 659 TGGGATGGGAGGCTGGCGAGACAGCGGGGAGAGCCATCCCTTACCCGAGA 718
 QY 183 TyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAspLysTyr 202
 Db 719 TACGTGACATCATCAACACAGCTGCGCTCAGATGTAGTATGACAGGAGGAGCTG 778
 QY 203 TrrPargLysPrrGlyValAsnGlyValAspLysTyrAspLysTrrArgGlyGlnLeu 222
 Db 779 TGGAGCTCTATGTACAGACACCATCCCTGGAG----- 811
 QY 223 IleGluAspValGlnHisThrPheGluGlnIleLysProLeuTyrGlnHisLeuHisAla 242
 Db 812 ---CAAGACCTGAGGGGCTCTCCAGAGCTGCGACCTCACTCACTCACTCACTCACTG 868
 QY 243 TyrValArgAlaLysLeuMetAsnAlaTyr---ProSerTyrIleSerProIleGlyCys 261
 Db 869 TAGGTGGCGGGGCGCTGCACCGCTACCTAGGGGCGCCAGACATCACTGAGGGGCGCC 928
 QY 262 LeuProAlaHisLeuLeuLysAspMetTrrPgluArgPheTrrPheAsnLeuTyrSerLeu 281
 Db 929 ATTCTGCTCAGCTGCTGGGAGACATGTGGGGCGAGCTGCTGCAACATCTATGACTTG 988
 QY 282 ThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGlnAla 301
 Db 989 GTGGTCCCTTCCTTCACGCCCTCGATGAGACACACAGAGCTATGCTAAGCAGCGCC 1048
 QY 302 TrpAspAlaGlnArgIlePheLysGluAlaGlyLysPhePheValSerValGlyLeuPro 321
 Db 1049 TGGACGCCACAGAGAGATTTAAGAGCGCTGATGATTTCTACCTCCCTGGGGCTGCTG 1108
 QY 322 AsnMetThrGlnGlyPheTrrPgluAsnSerMetLeuThrAspProGlyAsnValGlnLys 341
 Db 1109 CCCGTGCTCTGAGTTCTGGACACAGTCACTGCTGAGAGACCAACAGCGGGCGGAG 1168
 QY 342 AlaValCysHisProThrAlaTrrPaspLeuGlyLysGly---AspPheArgIleLeuMet 360
 Db 1169 GTGCTCTGCAACGGCTGCGCTGAGACTTCTACAAACGCAAGAGCTTCGGATCAAGCAG 1228
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyHisIleGln 380
 Db 1229 TGCACACCGGTGAATCTGGAGACCTGGTGTGGCCACCAAGAAATGGGCCCATCCAG 1288
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyLysPhe 400
 Db 1289 TATTTCATGACAGTAAAGACTTACCTGTGCTTGAAGGAGGCTGCCAACCCCGGCTTC 1348
 QY 401 HisGluAlaValGlyGlnIleMetSerLeuSerAlaIleThrProLysHisLeuLysSer 420
 Db 1349 CATGAGGCGCATTTGGGAGCTGCTACCTCTCAAGTGTATGACGCCAAGCACCTGACAGT 1408
 QY 421 IleGlyLeuLeuSerProAspPheGlnIleLysAsnGlnIleThrGlnIleAsnPheLeu 440

Alignment Scores:		
Pred. No.:	3,51e-147	4020
Score:	1337.00	255
Percent Similarity:	61.05%	118
Best local Similarity:	41.73%	204
Query Match:	31.16%	34
DB:	4	9
US-09-978-385-2 (1-805)	x	US-09-050-159-130 (1-4020)

[illegible]


```

QY 331 SerMetLeuThrAspProGluYasnValGlnLysAlaValCysHisProThrAlaTrpAsp 350
    |||||
Db 2852 TCATGCTGGAGAACCAACCGGCGGGGGTGTGGCCGCGCTGGAGC 2911
QY 351 LeuGlyLysGly--AspPheArgIleLeuMetCysThrLysValThrMetAspAspPhe 369
    |||
Db 2912 TTCTACAAACGGCAAGACTCCGATATCAAGACAGTGCACACCGTGAACCTGGAGACCTG 2971
QY 370 LeuThrAlaHisGlnMetGlyHisIleGlnIleTyrAspMetAlaTyrAlaIleGlnPro 389
    :|||
Db 2972 GTGGTGGCCCAACAGCAAGATGGCCACATCCATATTTCTGACAGTACAAAGCTTACCT 3031
QY 390 PheLeuLeuArgAsnGlnLysAlaAsnGlnLysPheHisGlnAlaValGluIleMetSer 409
    |||||
Db 3032 GTGGCCTTGAGGAGGGGTGGCAACCCCGCTTCATGAGCCATTTGGGGAGCTGTGACCC 3091
QY 410 LeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGln 429
    |||||
Db 3092 CTCTCAGTGTCTACGGCCCAAGCACTGCACAGTCTCAACCTGCTGAGCAGTGGGGTGGC 3151
QY 430 GluAspAsnGluThrGlnIleAsnPheLeuLysGlnAlaLeuThrIleValGlyThr 449
    |||
Db 3152 AGGAC--GAGCATGACATCACTTTCTGATGAAGATGGCCCTTGACAAAGTCCCTTT 3208
QY 450 LeuProPheThrTyrMetLeuGlnLysTrpArgTrpMetValPheLysGlyGluIlePro 469
    :|||
Db 3209 ATCCCTTGAGTACCTCTGCATGATGAGTGGCGCTGAGGGTATTGATGGAAGCATCC 3268
QY 470 LysAspGlnTrpMetLysLysTrpIleMetLysArgGluIleValGlyValGlu 489
    |||||
Db 3269 AAGGAACATATTAACAGGAGTGGTGGAGCCCTGAGCTGAAGTACAGGAGCCCTGCCCC 3328
QY 490 ProValProHisAspGluThrTyrCysAspProAlaSerIlePheHisValSerAsnAsp 509
    |||||
Db 3329 CCAGTCCCAAGACATCAAGTACCTTGAACCCAGGCGCAAGTTCACATTCCTTCAAGC 3388
QY 510 TyrSerPheIleArgTyrThrThrArgThrLeuTyrGlnPheGlnPheGlnGluAla 529
    :|||
Db 3389 GTGCCCTTACATGAGTACTTGTCTCAGCTTCATCCAGTTCAGTTCACAGGAGCACTG 3448
QY 530 CysGlnAlaAlaLysHisGlnLysProLeuHisLysCysAspIleSerAsnSerThrGlu 549
    |||||
Db 3449 TGCAGAGCAGCTGGCCACACAGGCGCCCTGCAACAGTGTGACATCTACCACTCAAGAG 3508
QY 550 AlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProTyrThrLeuAla 569
    |||||
Db 3509 GCCGGAGACGCGCTGGCGACGCCCATGAAGCTGGGCTTCAGTACGCGCTGGCGGAGGCC 3568
QY 570 LeuGluAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGlu 589
    :|||
Db 3569 ATGCAGCTGATACGGGCGACGCCCAACATGAGCGCTCGGCAATGTGAGTACTTCAAG 3628
QY 590 ProLeuPheThrTrpLeuLysAspLysAsnLys-----AsnSerPheValGlyTrp--- 606
    |||||
Db 3629 CCGCTGTGAGTGGCTCCGACGAGAGAACAGCTGCATGGGAGAGAGCTGGGCTGGCCG 3688
QY 607 SerThrAspTrpSerProTyrAlaAspGlnSer 617
    :|||
Db 3689 CAGTACAACTGACCGCGAAGCTCCGCTCGTCA 3721

```

```

RESULT 7
US-09-162-484-19
; Sequence 19, Application US/09162484
; Patent No. 6248724
; GENERAL INFORMATION:
; APPLICANT: Phillips, M. Ian
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS TARGETED TO
; FILE REFERENCE: UPLA/087/UPLA087P
; CURRENT APPLICATION NUMBER: US/09/162,484
; CURRENT FILING DATE: 1998-09-25

```

```

; EARLIER APPLICATION NUMBER: 60/059,661
; EARLIER FILING DATE: 1997-09-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 3942
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-162-484-19

Alignment Scores:
Pred. No.: 5,29e-144 Length: 3942
Score: 1310.00 Matches: 251
Percent Similarity: 61.20% Conservative: 115
Best Local Similarity: 41.97% Mismatches: 214
Query Match: 30.53% Indels: 18
DB: 4 Gaps: 7

US-09-978-385-2 (1-805) x US-09-162-484-19 (1-3942)

QY 20 ThrIleGlnGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeu 39
    |||
Db 1948 ACTGATGAAGCCCAAGGCTTAACAGTTTCGAGAGAGATGACCGGACAGCCAGAGTGTG 2007
QY 40 PheTyrGlnSerSerIleuAlaSerTrpAsnTyrAsnIleThrGluGluAsnVal 59
    :|||
Db 2008 TGGAAAGATTAAGCAGAGGCAAGCTGCATTTAAACACCAACATTAACATGAGGAGCAGC 2067
QY 60 GluAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeu 79
    |||
Db 2068 AAGATCTCTCTCAGAAAAAACAAGAGTGTCCAAACATTAACCTGAATATGAGCAGCTGG 2127
QY 80 AlaGlnMetTyrProLeuGlnIleGlnAsnLeuThrValLysLeuGlnAla 99
    |||||
Db 2128 GCCAAGACATTTGAGCTGAGCAACTTCCAGAACTTACCATCAAGCGGATCATTAAGAG 2187
QY 100 LeuGlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThrIle 119
    :|||
Db 2188 GTTCAGAACGTGACCGGCGGAGTGTGCTGCCCAACGAGTTAAGAAAGTCAACACAGATC 2247
QY 120 LeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGln 139
    |||||
Db 2248 CTGTAAGACATGAGAGACGACTTACAGTGTACCAATGTTTCTACAAAGTGC----- 2301
QY 140 GluCysLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsn 159
    |||||
Db 2302 ACTGTCTGTCTACGTGAGCTGATCTGACAAATGATGAGCCAGCTCCGAAATAGCAAG 2361
QY 160 GluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeu 179
    |||||
Db 2362 GAATGCTTGGGTGTGGAGAGAGCTGGCGAGACAAAGTGGGAGAGCCATCTTCCCTTT 2421
QY 180 TyrGlnGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAspTyr 199
    :|||
Db 2422 TTCCCAAAAGTACGTGAGCTTCCCAACAAAGATCCCAAGTCAAGGCTACTCTGATGCA 2481
QY 200 GlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArg 219
    |||||
Db 2482 GGGGATTCCTGGAAGTCTCTCATATGAGTCCGATGACTTGGAG----- 2523
QY 220 GlyIleuIleGluAspValGlnHisThrPheGlnGluIleLysProLeuTyrGlnHis 239
    :|||
Db 2524 -----CAAGACCTGGAAAACTATACAGAGAGCTGACGCCCTTACCTGAAC 2571
QY 240 LeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSer---TyrIleSerPro 258
    |||||
Db 2572 CTGATGCTTAATGTCGCCCGCTCCCTGACCGCATTAATGCTGAGTACATCAACCTG 2631
QY 259 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpLysArgPheThrIleAsnLeu 278
    |||
Db 2632 GATGGTCCCATTCCTGCCACACCTGTGAGAAATGTGGGACAGACTTGGTGGTCAACATC 2691
QY 279 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 298

```

```

Db 2692 TATGACTTGGTGCACCCCTTCCCTCCGCCCAAGTATATATCCAGGAGGCGCATGATA 2751
Oy AspglnalatrPaspAlaGlnargIlePheIysGluAlaIubIysPhePheValSerVal 318
Db 2752 AAGCAGGAGATGACACCAGAGAGATATTAAAGAACCTGACATATTTTACCCTCGTG 2811
Oy 319 GlyLeuProAsnMetThrGlnGlyPheThrGluAsnSerMetLeuThrAspProGlyAsn 338
Db 2812 GGGCTGTACCTGCTGCCCTGAGTCTGGAAAGTCAATGGTTAGAGAAAGCAACCCGAT 2871
Oy 339 ValGlnIysAlaValCysHisProThrAlatrPaspLeuIysGlySerVal 357
Db 2872 GGGAGGAGGTGGTGTGGCATGCTCCAGCCCTGGAGACTTCACAAAGGCAAGGACTTCAG 2931
Oy 358 IleLeuMetCysThrIysValThrMetAspAspPheLeuThrAlaHisIshIsgluMetGly 377
Db 2932 ATCAGAGCAATGACTCTGTCGACATGAGGAAATGGTATAGCCACCAAGAAATGGGC 2991
Oy 378 HisIleGlnIyrAspMetAlaIyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsn 397
Db 2992 CACATCCAGATATTTCATCAGTACAGAACACTTGGCTGTGACCTTTCGGGAGGGCGCAAC 3051
Oy 398 GluGlyPheHisIsgluAlaValGlyIleMetSerLeuSerAlaAlaThrProIysHis 417
Db 3052 CCCGCTTTCATGAGGCTATGGAGATGTTGGCTGTCTGTCTGTACACCCCAAGCAT 3111
Oy 418 LeuIysSerIleGlyLeuLeuSerProAspPheGlnIguAsnGluThrGluIleAsn 437
Db 3112 CTACACAGCTCAGCTCCTGAGCAGTGTAG--GGCAGTGGCTACGAGCATGACATCAGC 3168
Oy 438 PheLeuLeuIysGlnAlaLeuThrIleValGlyThrLeuProPheThrIyrMetLeuGlu 457
Db 3169 TTTCTAATGAATGAGCCCTTGACAAATTCGCTTCATCCCTTCAGTACCTATGATGAC 3228
Oy 458 LysTrpArgTrpMetValPheIysGluIleProIysAspGlnTrpMetIysLysTrp 477
Db 3229 CAGTGGCGGTGAGGCTCTTTTACGGAAGCATCACAGAGAACTACAAACAGAGATGG 3288
Oy 478 TrpGluMetIysArgGluIleValGlyValGluProValProHisAspGluThrIyr 497
Db 3289 TGGAGCTCTCAGACTGAAAGTACAGGAGGCTGCTCCCTCCAGTGGCTTACCAAGGTGAC 3348
Oy 498 CysAspProAlaSerLeuPheHisIysValSerAsnAspTrpSerPheIleArgTrpThr 517
Db 3349 TTGAGCCAGGCTCAGTCCAGTCTCTGCGAATGTCGCAATATAGTACTTATATC 3408
Oy 518 ArgThrLeuTrpGlnPheGlnPheGlnIguAlaLeuCysGlnAlaAlaIysHisIsgluGly 537
Db 3409 AGCTTCATCATCCAGTTCAGTTCACAGGAGCATATGTGCGGCGAGCGGACACACCGGC 3468
Oy 538 ProLeuHisLysCysAspIleSerAsnSerThrGluAlaGlyIguIysLeuPheAsnMet 557
Db 3469 CCCCTGTACAGATGATATCTTACCAATCCAAAGAGAGGAGGAGGAGCTCTGCGAGATGCC 3528
Oy 558 LeuArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLys 577
Db 3529 ATGAAGTGGGCTACAGTACAGAGAGGCGCAAGAGCATATATCAGAGCAACCT 3588
Oy 578 AsnMetAsnValArgProLeuLeuAsnTrpPheGluIguProLeuPheThrIleLeuLysAsp 597
Db 3589 AACATGTACAGCTCTGCTCATTTATGATTTACTTACAGCCACTGACGAAATGGCTCTGCACA 3648
Oy 598 GluAsnLys-----AsnSerPheValIguIyr-----SerThrAspTrpSerPro 612
Db 3649 GAGAACGAGGAGACATGAGAGACATGGCGCTGGCGGAGTACACTGGACACCA 3702

```

```

; TITLE OF INVENTION: Zacl: A Human Metalloenzyme
; FILE REFERENCE: 98-79
; CURRENT APPLICATION NUMBER: US/09/440,325A
; CURRENT FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2082
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate sequence encodes the amino acid
; OTHER INFORMATION: sequence of SEQ ID NO:1.
; NAME/KEY: misc-feature
; LOCATION: (1)-(2082)
; OTHER INFORMATION: n = A,T,C or G
US-09-440-325A-2

Alignment Scores:
Pred. No.: 7,86e-75 Length: 2082
Score: 721.00 Matches: 174
Percent Similarity: 43.47% Conservative: 79
Best Local Similarity: 29.90% Mismatches: 299
Query Match: 16.80% Indels: 30
DB: 4 Gaps: 14

US-09-978-385-2 (1-805) x US-09-440-325A-2 (1-2082)
Oy 22 GluGluGlnAlaIysThrPheLeuAspLysPheHisIsgluAlaGluAspLeuPheThr 41
Db 157 GACAGCAGACCAATATTTTTCATTTTATATATATATATATATATATATATATATATAT 216
Oy 42 GlnSerSerLeuAlaSerTrpAsnTrpAsnTrpAsnIleThrGluGluAsnValGlnAsn 61
Db 217 AARTTATGAGCAGCAGACATGGAATATATATATATATATATATATATATATATATAT 276
Oy 62 MetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrIleAlaGln 81
Db 277 ATGATGAAGAGATGATGAGAGG--WSNCARTTATGATTTATTTTGGAGACACGACAY 333
Oy 82 MetIyrProLeuGlnGluIleGluAsnLeuThrValIysLeuGlnIleGlnAlaLeuGln 101
Db 334 YNTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 393
Oy 102 GlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeuAsn 121
Db 394 AAVATGAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 453
Oy 122 ThrMetSerThrIleIysSerThrGlyIysValCysAsnProAspAsnProGluCys 141
Db 454 NNNTYTGARATGACNTATWSNATGGCNCARGTNGTYTNAAYGARGGCCN-----TGY 507
Oy 142 LeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTrpAsnGluArg 161
Db 508 YTWMSYNTNGAR---WSNCARTTNGARTTATGCGACNWSNMCGAAYAAARGARGARTN 564
Oy 162 LeuTrpAlaTrpGluSerTrpArgSerGluValGlyIysGlnLeuArgProLeuArgGlu 181
Db 565 YTNVGGCGNTGGCARGNGTGCARGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 624
Oy 182 GluTrpValValLeuLysAsnGluMetAlaArgAlaAsnHis-----TyrGluAspTrpGly 200
Db 625 CAYTATGTGARTTWSNATATATATATATATATATATATATATATATATATATATATAT 684
Oy 201 AspTrpTrpArgGlyAspTrpGluValAsnGlyValAspGlyTrpAspTrpSerArgGly 220
Db 685 GCYVNTMGCAVWSNATATATATATATATATATATATATATATATATATATATATAT 723
Oy 221 GlnLeuIleGluAspValAlaHisThrPheGluGluIleLysProLeuTrpGluHisLeu 240
Db 724 -----CARGAYTNGARGNGNTTTCARGARARNTNMGCCYTTTATTTAAATCCN 774
Oy 241 HisAlaIyrValArgAlaLysLeuMetAlaAsnAlaIyr-----ProSerTrpIleSerProIle 259

```

```

RESULT 8
US-09-440-325A-2
; Sequence 2, Application US/09440325A
; Patent No. 6280994
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.

```

```

Db 775 CAYACNTAYGTNMGNGMNCNTNCAVMGNCATATAGCCNGARVYTNATHGAYTYTNMGN 834
QY 260 GLYCYSEUPROALAHISLEULEUGLYASP---METIRPGLYARGPHEITRPHASPLEU 278
Db 835 GGNCNCNATGCCNCCNCAATYTNMGNGARAAVACNNTYTGCCNCAWSTGGTNAATAVTH 894
QY 279 TYRSELEUTHRALPROPHEGLYGLNLSYPROASNILEASPVALTTHASPALAMEVAL 298
Db 895 YTMGAVCCNCTNTYTNCCNTTYTNNAARARATGCCNGARAGVGNACMAAATHTAGAR 954
QY 299 ASPGLALATIRPASPALAGLARG---ILEPHELYGUALAGLULYSPEPHEVALSER 317
Db 955 GTWCACACATYTGGAARCCNARARARYTNATGYTNARGARCGNARACNTTYTTACTAY 1014
QY 318 VALGLY-----LEUPROASNETHRGINDLYPHERTPGLUASNSEMETLEUTHRASP 335
Db 1015 YTMGNGVYTNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCN 1074
QY 336 PROGLYASPVALGLNLYALVALCYSHISPROTHRALATIRPASPGLULYS---GLY 354
Db 1075 CCNACNGATGNGMNGANGARTGYCAYATHMSNGCNTGGAATYTTAYCAKAGAYGAY 1134
QY 355 ASPHEARTGILEUMECSYTHRLYSVALTHMETLASPASPHELEUTHRALAHISHS 374
Db 1135 GAYTYTMGNATHAARARTGCGNGARGTNACNACNGARAYCCNTNMSNATHTTYCAY 1194
QY 375 GLUWETGLYHISILEGINTYRASPMEVALATYRALAGLAPROPHASELEUARGSN 394
Db 1195 CARATGGGNCAYTTYCARATYTYTYTCARTAYAAARAYTYMNSNATHTHTTYGNACN 1254
QY 395 GLYLAASNGIUGLYPHEHISGLUALVALGLYGLULIEMESERLEUSERALALATHR 414
Db 1255 GGNCNCNAYCCNCCNTTYGARARCGNCGNNGMNGTATACNATYTNMSNCCNCCNCCN 1314
QY 415 PROLYSHISLEUSERIIEGLYLEULEUSERPROASPHEGLNGLUASPAGLULTR 434
Db 1315 CAYARCAVYTYTNAAVATHTGNTYTNMSN-----YTNVTMGARCAV 1359
QY 435 GLULIENASPHLEULEUSGLNALALEUTHRLILEVALGLYTHLEUPROPHETHR 454
Db 1360 GAROTNAAYTYTYTNATGCAVATHTGNTYTNARARATHTGNTYTAHCCNTTYGNTAY 1419
QY 455 METLEUGLULYSIRPARGTRPMETVALPHELYSGLYGLULIETPROLYSASPLTIRMET 474
Db 1420 YTNATGCAVYTYTNMGNTGGAARNTTYGAYGNGNACNATHTGGAARAGATHTATAAY 1479
QY 475 LYSLESTIRTPGLU---METLSARGGLULILEVALGLYVALGLUPROVALPROHIS 493
Db 1480 CARGARTGSGGAAYTYTMNGMNGANTNARATAYCARGNGNTNTGCGCCNCCNATHTCCNCAV 1539
QY 494 ASPGLUTHTYRCYASPROALASERLEUPHEHISVALSERASNPASYRISERPHELLE 513
Db 1540 WMSGARGARAGATYTYGACCGNGCNGNARARTTYCATYTYMNSCNGCNGCNGCNGCNGCNG 1599
QY 514 ---ARGTYTYRTHARGTHLEUTHRGINDLYPHEGLNGLULIEMESERLEUSERALALATHR 1599
Db 1600 MGNMNTAYTYTYTNMSTYTNMGNTYTNATGCAVATHTATAAYNAARATHTCCNCCNCA 1659
QY 533 ALALYSHISGLIUGLYPROLEUHSILYSYASPLISERASNPASERHISGLUALAGLULIN 552
Db 1660 WMSGNCNAYATGNGCNCNTNCAVARTGYATHTATAAYNAARATHTCCNCCNCA 1719
QY 553 LYSLEUPHEASNEMETLEUARGLEUGLYLSERGLUPROTPRTHLEUALALEUGLUSN 572
Db 1720 -----YTNNTNCGNTYTNARATYTNMGNSWMSNARCCMTGGCCNGARGTNTNARATG 1773
QY 573 VALVALGLYALALYSASNEMETLASVALARGPROLEULEASNLYRPHGLUPROLEUPHE 1773
Db 1774 YTNACNGNARMSNGARGTNMSNACNAYGNTTYATGATACNTATYTAARCCNTYNTN 1833
QY 593 THRTIRP 594
|||||

```

```

Db 1834 ACNTGG 1839
RESULT 9
US-08-664-596B-14
; Sequence 14, Application US/08664596B
; Patent No. 5607703
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallee, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; APPLICANT: Spaulding, Ylki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/664,596B
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELEPHONE: (617) 498-9224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEO ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-664-596B-14
Alignment Scores:
Pred. No.: 3 88e-46
Score: 469.00
Percent Similarity: 98.96%
Best Local Similarity: 98.96%
Query Match: 10.93%
DB: 1
Gaps: 0
US-09-978-385-2 (1-805) x US-08-664-596B-14 (1-467)
QY 1 MetSetSerSetSerTIRPleuleuSerleuValAlaValThralAlaGlnSerThr 20
Db 182 ATGTGAAGCTCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 241
QY 21 ILEGLUGLINALALYSYTHRPHLEUASPLYSPEASNHISGLUALAGLULASPLEUPHE 40
Db 242 ATTTGAAGAACGACACATTTTGGACAAGTTTAACACGACGACGACGACGACGACGACGAC 301
QY 41 TYRGLNISRSELEUALA-SERTIRPASNLYRASNTHIRASNTHIRHISGLULIASNVALGI 60
Db 302 TATCAAGTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
QY 60 nasmetasnasnalaglaspyspyspyspyspyspyspyspyspyspyspyspyspyspys 80

```

```

Db      362 AANATGATATGCTGGGACAAATGCTCTTTTAAAGGACGTCACACTTGC 421
      |||
Qy      80 aglmetcyrProleugInclutleGlnAsnleuThrValIleu 95
      |||
Db      422 CCAATGATCCACTACAGAAATTCAGATCTCACAGTCAGCTT 467
      |||

RESULT 10
US-08-905-223-27
; Sequence 27, Application US/08905223
; Patent No. 622029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Ducleret, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 848 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; DEVELOPMENTAL STAGE: Fetal
; TISSUE TYPE: Kidney
; FEATURE:
; NAME/KEY: sig-peptide
; LOCATION: 32..73
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 10.7
; OTHER INFORMATION: seq IMLLFLVTAIHA/EL

US-08-905-223-27

Alignment Scores:
Pred. No.: 2.72e-35 Length: 848
Score: 381.00 Matches: 87
Percent similarity: 58.54% Conservative: 33
Best Local Similarity: 42.44% Mismatches: 63
Query Match: 8.88% Indels: 22
DB: 4 Gaps: 5

US-09-978-385-2 (1-805) x US-08-905-223-27 (1-848)
Qy      612 ProTyrAlaSerIleuValArgIleSerLeuYsSerAlaLeuGlyAspLys 631
      |||
Db      86 CCAAGTCACAGAAATGCTTTTAAAGTACGACTAGTACAGAACAGCTCTGGGAGATPAA 145
      |||

```

```

Qy      632 AlaTyrGluTrpAsnAspAsnGluMetTyrLeuPheArgSerSerValAlaTyrLamMet 651
      |||
Db      146 GCATATGCTGGTATACCAATGAGAAATACCTTTAAACCGATGGTACTTCCACTC 205
      |||
Qy      652 ArgGluTyrPheLeuYsValIleYsAsnGluMetIleuPheGlyGluAspValArg 671
      |||
Db      206 AGA-----AAAGTCCACACAGAGAACAAACA---GAAATTTCCATGCTCTTA 250
      |||
Qy      672 ValAlaAsnLeuYsProArgIleSerPheAsnIlePheValThrAlaProLysAsnVal 691
      |||
Db      251 CTTTGCAATGTATACCCAGAGGGGTATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 307
      |||
Qy      692 SerAspIleIleProArgThrGluValGluYsAlaIleArgMetSerArgSerArgIle 711
      |||
Db      308 AATCACACCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367
      |||
Qy      712 AsnAspAlaPheArgLeuAsnAspAsnSerLeuGluIlePheGluYsIleGlnProThrLeu 731
      |||
Db      368 AACATGCGCTCTTCTTAATGACCAACCTGGAATTTTAAATAATCCCTTCCACACTT 427
      |||
Qy      732 GlyProProAsnGluProProValSerIleThrPheIleValPheGlyValIleMetGly 751
      |||
Db      428 GCACACACCCTGAGCCCATCTGTGCCCATCTGTGATTTATTTGTTGATATTTTGC 487
      |||
Qy      752 ValIleValValGlyIleValIleLeuIlePheThrGlyIleArgAspArgLysLys 771
      |||
Db      488 ATCATCATGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 547
      |||
Qy      772 AsnLysAlaArgSerGly----- 777
      |||
Db      548 AACAAAGAACCATCTGACATGATGACGCTGAARATTAATGTGAACATGATCAAT 607
      |||
Qy      778 GluAsn-----ProTyrAlaSerIleAspIleSerLysGlyGluAsnAsnProGlyPhe 795
      |||
Db      608 GAAATGCAATCCCTCTGATCCCTGACATGACATGACATGACATGACATGACATGACATG 667
      |||
Qy      796 GluAsnThrAspAsp 800
      |||
Db      668 ATGACAGAGATGAG 682
      |||

RESULT 11
US-09-247-155-27
; Sequence 27, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Ducleret, Aymeric
; APPLICANT: Bouquelere, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; EARLIER FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099,273
; EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 27
; LENGTH: 848
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: sig-peptide
; LOCATION: 32..73
; OTHER INFORMATION: Von Heijne matrix

US-09-247-155-27

Alignment Scores:

```

Pred. No.: 2.72e-35 Length: 848
 Score: 381.00 Matches: 87
 Percent Similarity: 58.54% Conservative: 33
 Best Local Similarity: 42.44% Mismatches: 63
 Query Match: 8.88% Indels: 22
 DB: 4 Gaps: 5

US-09-978-385-2 (1-805) x US-09-247-155-27 (1-848)

```

QY 612 ProTyrAlaAspGlnSerIleValAlrGleSerIleuLysSerAlaLeuGlyAspLys 631
DB 86 CGAGGTGCGAGAAATGCTTTTAACTGAGACTTACTATCAGAACAGCTCTGGAGATATA 145
QY 632 AlaTyrGluTyrPasnAspAsnGluMetTyrIleuPheAspGlySerValAlaTyrAlaMet 651
DB 146 GCATATGCTGGATACCAATCAATAAATCTCTCAAGAGAGGTGCTTCTCCATG 205
QY 652 ArgGlnTyrPheLeuLysValLysAsnGluMetIleuPheGlyGluGluAspValArg 671
DB 206 AGA-----AAAGTTCACACAGAGAGACACA---GAAATTTCCCATGTCCTA 250
QY 672 ValAlaAsnLeuLysProArgIleSerPheAsnPhePheValThrAlaProLysAsnVal 691
DB 251 CTTTGCATATGTAACCCAGAGAGGTATCATCTGTGTGTACAGACCTTCAAAA--- 307
QY 692 SerAspIleIleProArgThrGluValGluLysAlaIleArgMetSerArgSerArgIle 711
DB 308 AATCACCCTCTTCCGCTGCTGTGAGGTGCATACAGCCATTAAGAAAGAACAGACCGGATC 367
QY 712 AsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLeuGlyIleGlnProThrLeu 731
DB 368 AACATGCTCTTCTTAAATGACCAAACTCTGGAATTTTAAAAATCCCTCCACACTT 427
QY 732 GlyProProAsnGlnProProValSerIleThrIleuIleValPheGlyValMetGly 751
DB 428 GCACACCACCATGAGCCCTGCTGTCGCCATCTGATATATATATATGCTATTTTTCG 487
QY 752 ValIleValAlaGlyIleValIleValIleuPheThrGlyIleArgAspArgLysLys 771
DB 488 ATCATCATAGTTGCATATGCACTACTGATTTTATCAGAGGATCTGCAACGTAADARAAG 547
QY 772 AsnLysAlaArgSerGly----- 777
DB 548 AACAAAGAACCATCTGAAGTGATGAGCCGTGAARATAKTGTGAACATCATGATCAATT 607
QY 778 GluAsn-----ProTyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPhe 795
DB 608 GAAATGCGATCCCTCTGATCCCTGACATGGAAGGAGGACATATTAATGATGCTTC 667
QY 796 GluAsnThrAspAsp 800
DB 668 ATGACAGAGATGAG 682

```

RESULT 12

US-08-157-171-6

; Sequence 6, Application US/08157171

; Patent No. 5736323

; GENERAL INFORMATION:

; APPLICANT: Soubrier, Florent

; APPLICANT: Hubert, Christine

; APPLICANT: Corvol, Pierre

; TITLE OF INVENTION: Agents and Procedures for the Study of

; TITLE OF INVENTION: the Genetic Polymorphism of the Angiotensin I Converting

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Merchant & Gould

; STREET: 90 South 7th Street, 3100 No. 5736323west Center

; CITY: Minneapolis

; STATE: MN

; COUNTRY: USA

; ZIP: 55402

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/157,171
 FILING DATE: 24-JAN-1994
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Hillson, Randall A.

REGISTRATION NUMBER: 31,838

REFERENCE/DOCKET NUMBER: 8076.1030USNO

TELEPHONE: 612-332-5300

TELEFAX: 612-332-9081

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 192 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

IMMEDIATE SOURCE:

CLONE: Exon 18

US-08-157-171-6

Alignment Scores:

Pred. No.: 4e-07 Length: 192

Score: 133.50 Matches: 30

Percent Similarity: 56.16% Conservative: 11

Best Local Similarity: 41.10% Mismatches: 21

Query Match: 3.11% Indels: 11

DB: 1 Gaps: 2

US-09-978-385-2 (1-805) x US-08-157-171-6 (1-192)

```

QY 196 TyrGluAspTyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyr 215
DB 3 TATGATGATGACAGGAGCTGCTGAGGTCTATGTACGACACCATCCCTGGAG----- 56
QY 216 AspTyrSerArgGlyLeuIleuIleGluAspValGluHisThrPheGluGluIleLysPro 235
DB 57 -----CAAGACTGTGAGCGGCTCTCCAGAGCTGACAGCA 92
QY 236 LeuTyrGluHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyr---Proser 254
DB 93 CTCTACCTCAACCTGATGCTGACGTGCGCGGCGCTGCACCGTCACTACGAGGCGCCAG 152
QY 255 TyrIleSerProIleGlyLysLeuProAlaHisLeuLeu 267
DB 153 CACATCAACCTGAGGAGGCGCATTCCTGCTCACTGCTG 191

```

RESULT 13

US-08-392-625-16

; Sequence 16, Application US/08392625

; Patent No. 5837485

; GENERAL INFORMATION:

; APPLICANT: Entian, Karl-Dieter

; APPLICANT: Gtz, Friedrich

; APPLICANT: Schnell, No. 5837485bert

; APPLICANT: Augustin, Johannes

; APPLICANT: Engelke, Gerhard

; APPLICANT: Rosenstein, Ralf

; APPLICANT: Kaletta, Cortina

; APPLICANT: Kleinf, Cora

; APPLICANT: Wieland, Bernd

; APPLICANT: Kupke, Thomas

; APPLICANT: Jung, G nther

; APPLICANT: Kellner, Roland

; TITLE OF INVENTION: Biosynthetic Process For The Preparation

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 STREET: 1100 New York Avenue
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 FILING DATE: US/08/392.625
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/876,791
 FILING DATE: 30-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Esmond, Robert W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 0652, 0980002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8700 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-392-625-16

Alignment Scores:
 Pred. No.: 0.00153 Length: 8700
 Score: 127.00 Matches: 152
 Percent Similarity: 32.58% Conservative: 121
 Best Local Similarity: 18.14% Mismatches: 261
 Query Match: 2.96% Indels: 304
 DB: Gaps: 39

US-09-978-385-2 (1-805) x US-08-392-625-16 (1-8700)

QY 1 MetSerSerSerSerTripleuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
 DB 2562 ATTAGTGAAGCAGCATATATCTCTGTTATATCTCTATCATTTGGTACAAAACCT 2621
 QY 21 lIeGluglInAlaIysThrPheLeuAspLysPheAsnHISglu- 35
 DB 2622 ATTAGCAATTATCAGCAATTTTATGATATAATGATTTGAACAACAGTAATTTTA 2681
 QY 36 -----AlaGluAspLeu-----PheTyrGlnSerSerLeuAlaSerTyr 50
 DB 2682 AAGCAATTCCTCTCAGATATAATGATTTGCTATCCCAAAAAAGACAGTTATAGTTT 2741
 QY 51 AsnThrAsnIleThrGluGluAsnValGlnAsnMetAsnAlaGlyAspLysTyrSer 70
 DB 2742 TCTAATTAACATT----- 2753
 QY 71 AlaPheLeuLysGluGlnSerThrLeuAlaGlnMetTyrProLeuGlnGluIleGlnAsn 90
 DB 2754 GCATTTTAAAGAAAG----- 2771
 QY 91 LeuThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGlu 110
 DB 2772 -----TATTGCTTGCATTCATAAATACAGCCATTTGAATTAACAGAA 2816
 QY 111 AspLysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGly 130
 DB 2817 AACGACGTTAAATTTTAAAGAAAG--AATATACAGTTTCTAAATACATGCG----- 2867
 QY 131 LysValCysAsnProAspAsnProGlnGluLysLeuLeuLeuGluProGlyLeuAsnGlu 150

DB 2868 -----CCTGTTCAACTGAA 2882
 QY 151 lIeMetAlaAsnSerLeuAspTyrAsnGluAlaGluLeuThrPalaTyrGluSerTyrArgSer 170
 DB 2883 ATA-----TATAGTGAGATATATTT- 2903
 QY 171 GluValGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGlnMet 190
 DB 2904 -----GGAATATTCATTAAGGT-----TATGAGATTTTCCGCTGATAGTCAATATTA 2954
 QY 191 AlaArgAlaAsnHISLysGluAspTyrGlyAspTyrTyrArgGlyAspTyrGluValAsn 210
 DB 2955 GATCTTTTAAATCCGCGTGAACCTTTGGAGGTT--ACGGAAATTTCAATTAATAAG 3011
 QY 211 GlyValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGlnIsthThrPhe 230
 DB 3012 -----AAAAAATTCATTAACAAGAAATTAAGTGTGAT----- 3044
 QY 231 GluGluIleLysProLeuTyrGluHISLeuHISAlaTyrValArgAlaLysLeuMetAsn 250
 DB 3045 -----CATTAACATTAATTCATGAAATGAAGATGTTTGA 3080
 QY 251 AlaTyrProSerTyrIleSerProIleGlyCysLeuProAlaHISLeuGlyAspMet 270
 DB 3081 ATAGCCCAATTAATGAAGGTCCTTAACCTCAAGAAATGTAAATATTTTGAATAT- 3137
 QY 271 TrrGlyArgPheThrPheThrAsnLeuTyrSerLeuThrValProPheGlyGlnAspProAsn 290
 DB 3138 ---AATAGAAATATATATCTTTTAATTTAAATTTTAACT-----AAAAGTGA 3185
 QY 291 lIeAspValThrAspAlaMetValAspGlnAlaTyrAspAlaGlnArgIlePheLysGlu 310
 DB 3186 ATGATATTAATGACATATTTATTTGAGGCTACATTTAAACAACCTTATCTATTTCTGAA 3245
 QY 311 -----AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnGly 326
 DB 3246 AAGCATGATTCAGAAATTTGATTTGATTAATTCATTAATTAATTAATGATTTGATTT 3305
 QY 327 -----PheTyrGluAsnSerMetLeuThrAspProGly 337
 DB 3306 GAATTAACAATTTTAAAGAAATTTTCAATTGAAAAAACAATAATTTATACAACTTAA 3365
 QY 338 Asn-----ValGlnLysAlaVal 343
 DB 3366 ACTGAAAGAGCATTCATTCATTCCTTTTGTCAAGAAATTTATTAATAATATTAAT 3425
 QY 344 CysHISProThrAlaTyrAspLeuGlyLysGlyAspPheArgIleLeuMetCysThrLys 363
 DB 3426 TTAAGACCACTTCTGGAAATTAATTCGAATGTTTTCGAA-----ACTGAA 3476
 QY 364 ValThrMetAspAspPheLeuThrAlaHISLysGluMetGlyHISlIeGlnTyrAspMet 383
 DB 3477 AATGTTTAATATAGGTTCCGCAACTAT-AGAGAAATAATGCGATATTCCAAAAGATGTA 3533
 QY 384 AlaTyrAla----- 386
 DB 3534 ATTAATTCCTTTGGAGATTAATCGATTGCTATTAATTTATTAATGACAAAGATTCAT 3593
 QY 387 -----AlaGlnProPheLeuAsnArgAsnGlyAlaAsn 397
 DB 3594 ATACTAAAAAAGACCTAAAAAAGATGATGAGATTCGATATTTAGAAAGCTTATACAT 3653
 QY 398 GluGlyPheHISgluAlaValGlyGluIleMetSer----- 409
 DB 3654 GAATCTAATATATGAGAGAAATTTGTAAGATTTTACGCCATTTATATAAAAAACTAGTTTA 3713
 QY 410 -----LeuSerAlaAlaThrProLys-----HISLeuLys--- 419
 DB 3714 AAGAGCAATCTTTCAATTAATCAATTAATAATAGAAATAGCACTTCATTAATCTTAAGAT 3773
 QY 420 -----SerIleGlyLeuLeuSerProAspPheGlnGluLysAsnGluThrGluIleAsn 437
 DB 3774 TGGTTTCAATTCATTTAAGTATTCCTTAATAAACATACCAAGAT-----AAT 3818


```

QY 51 AsnThrAsnIleThrgInGluAsnValGlnAsnMetAsnAsnIaGlyAspLysTrpSer 70
Db 2742 TCTATATACATT----- 2753
QY 71 AlaPheLeuLysGluGlnSerThrLeuAlaGlnMetLysTrpLeuGlnIleGlnAsn 90
Db 2754 GCATTTTAAAGAAAG----- 2771
QY 91 LeuThrValLysLeuGlnIleuGlnAlaLeuGlnIleuGlnLysSerValLeuSerGlu 110
Db 2772 -----TATTGCTTGCAATTCAAAATACAGCCCATATTGAATACAGAA 2816
QY 111 AspLysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleLysSerThrGly 130
Db 2817 AACGACGTTAAATAATTAGAAAG--AATATACAGTTTCTAAATCAATGCG----- 2867
QY 131 LysValLysAsnProAspAsnProGlnGluCysLeuLeuLeuGlnLysProGlyLeuAsnGlu 150
Db 2868 -----CCTGTTTCACACTGAA 2882
QY 151 IleMetAlaAsnSerLeuAspLysArgLysGlnIleuArgLeuThrPheIleuSerTrpArgSer 170
Db 2883 ATA-----TATAGTACAGTATATTT----- 2903
QY 171 GluValGlyLysGlnLeuArgProLeuLysGlnIleuValValLeuLysAsnGlnMet 190
Db 2904 -----GGAAATTCATATAAAGGT--TATGAGCATTTTCCCGTGAATGCCAATATTA 2954
QY 191 AlaArgLysAsnHisLysTrpGluAspLysGlyAspLysTrpArgGlyAspLysGluValAsn 210
Db 2955 GGATCTTTTAAATGCCGTGCAACTTTTGGAAAGTTT--ACGGCAAAATTTCAATATTAAG 3011
QY 211 GlyValAspGlyLysAspLysArgLysGlnIleuLysLeuAspValGlnIleuHisThrPhe 230
Db 3012 -----AAAAAATCAATTACAAAAAGAAATAGTCAT----- 3044
QY 231 GlnGluIleLysProLeuLysGlnIleuHisLeuHisAlaLysValArgAlaLysLeuMetAsn 250
Db 3045 -----CATTCATATATATACATGAATGAATGAATGGTTTGA 3080
QY 251 AlaLysTrpSerLysLysSerProLysGlyCysLeuProAlaHisLeuLeuGlyAspMet 270
Db 3081 ATAAAGCAATTAATGAAGGTCCCTCTTAACCTCAAGAAATGAATATTTTGAATAT-- 3137
QY 271 TrpLysArgPheTrpThrAsnLeuLysSerLeuThrValProPheGlyGlnLysProAsn 290
Db 3138 ---AATGAAATATATATCTGTTTAATTTAAATTTACT-----AAAAGTGA 3185
QY 291 IleAspValThrAspAlaMetValAspGlnAlaIleTrpAspAlaGlnArgIlePheLysGlu 310
Db 3186 ATGATATTAATGACATATTTAATGAGCTACATTTAACAACCTTATCATATCTCGAA 3245
QY 311 -----AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnLys 326
Db 3246 AAACATGATTCACAGAAATGTGATTCGTAATCTAATCTAATGTTAATATGAGTTGGATCT 3305
QY 327 -----PheTrpGluAsnSerMetLeuThrAspProGly 337
Db 3306 GAATTTATACAAATTTTAAAGAGAAATTTCAATTGAAAAAACAATTTATACAACTATA 3365
QY 338 Asn-----ValGlnLysAlaVal 343
Db 3366 ACTGAAGAGCATGTGACTCATTAACCTTTTGTCCAAAGATTTATTTATATAAAATATATT 3425
QY 344 CysHisProThrAlaTrpAspLysGlyLysPheArgIleLeuMetCysThrLys 363
Db 3426 TTAACACCGACTTGGAAAAATTAATCGAAATGTTTCTGAA-----ACTGAA 3476
QY 364 ValThrMetAspAspPheLeuThrAlaHisGlnMetCysGlyHisIleGlnIleLysTrpMet 383
Db 3477 AATGGTTAAATAGTTTCGCACTATT--AGAGAAAAATGGCATTTTCCAAAAAGATGTA 3533
QY 384 AlaTrpAla----- 386
Db 3534 ATTATTCGCTTTGGAGATATATGCGATTGCTATTAATTTATTAATGACAGACTGCATT 3593
QY 387 -----AlaGlnProPheLeuLeuArgAsnGlyValAsn 397
Db 3594 ATACTAAAAAAGAACTAAAAAAGACGTGAGATTCGAAATGTAGAAAGCTTTTCAAT 3653
QY 398 GlnGlyPheHisGlnAlaValGlyLysIleMetSer----- 409
Db 3654 GAATCTAAATTAAGAGAAATGTAGAAATGTAGCCATTATTAATAAAAAACTAGTTTA 3713
QY 410 -----LeuSerAlaAlaThrProLys-----HisLeuLys--- 419
Db 3714 AAAGAACATCTTTCTACTTTACCTAAAAATAGAAATAGACACTGCATTAATCTTAAGAT 3773
QY 420 -----SerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluLeuAsn 437
Db 3774 TGGTTTCAATTCATTTAAGATTTCTTAACATACCAAGAT-----AAT 3818
QY 438 PheLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrLysMetLeuGlu 457
Db 3819 TTTATTCAGATTAATCTATTACCATTTTAACGGAATTAAGAACTTAATTAATTTATTAAT 3878
QY 458 LysTrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysTrp 477
Db 3879 AATTTTATTTACATATAAATTTAAAGAGAT-----GAAAGATTATATAA----- 3923
QY 478 TrpGluMetLysArgGluIleValGlyValGluProValProHisAspGluThrLys 497
Db 3924 TTAAGATTATTTAAGAA-----GATGAAGATTAAT 3953
QY 498 CysAspProAlaSerLeuPheHisValSerAsnAspLysSerPheIleArgTyr----- 516
Db 3954 -----TCTCAATTTATTTCTTCATATAAATAATGGGAA 3986
QY 517 -----ThrArgThrLeuLysGlnIlePheGlnPheGlnAlaLeuCysGln 531
Db 3987 GATTATGCTGTATTAAATAGTAGAATTAATATGACTTCTGATAGTTGATTATGCTCGAA 4046
QY 532 AlaAlaLysHisGlnGlyProLeuHisLysCysAspIleSerAsnSerThrGluAlaGly 551
Db 4047 GTATAGATATAGTGCGTCCACACGCTAATGGAATATTTGAATTTTATATGATAT 4106
QY 552 GlnLysPheAsnMetLeuArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGlu 571
Db 4107 ACTCATTTA-----TCAATTAATATTAATACATCAGAG--TTCAAAATTCCAAAAAGAA 4157
QY 572 AsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTrpPheGluProLeu 591
Db 4158 TTTATCGTTGCT-----ATATCATAAGATTTTATTAATTAATTAATTAATTAATTAAT 4202
QY 592 PheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGlyTrpSerThrAspTrpSer 611
Db 4202 ----- 4202
QY 612 ProTyrAlaAspGlnSerIleLysValArgIleSerLeuLysSerAlaLeuGlyAspLys 631
Db 4203 -----ATTAAATAAAGAGAGAAAGAAAGAAATTTTAATTAATAATAGCG-----GAAGAT 4250
QY 632 AlaTyrGluTrpAsnAspAsnGlnLysMetLysLeuPheArgSerSerValAlaTyrAlaMet 651
Db 4251 TTTATTCGTAAGTAC-----ATA 4271
QY 652 ArgGlnTrpPheLeuLysValLysAsnGlnMetIleLeuPheGlyGluGluAspValArg 671
Db 4272 AGAGAAATAT----- 4280
QY 672 ValAlaAsnLeuLysProArgIleSerPheAsnPhePheValThrAlaProLysAspVal 691
Db 4281 ---AAAAATTTATTAAGCTTAACCTT-----ACCAATTCCTAAAAAATGAC 4319
QY 692 SerAspIleIlePro----- 696

```

```

Db 4320 TATGAAATTTTAAAAAGAAATTCGAACTTCATGAAATTTCTATTAAATAATAGT 4379
      :::::
Qy 697 -----ArgThgluValGluValAlaIleArgMetSerArgSerArgIle 711
      :::::
Db 4380 ATTTAGAAATCTTAAAGACACATAAAAAAGCTTAATACTTACGCTTCAGGATA 4439
      :::::
Qy 712 AsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLeuGlyIleGlnPro 729
      :::::
Db 4440 ATTGGCACTTTT-----ATTAACATCGCTTGTAATAGAAATATTCGCTATTAAATCCT 4490
      :::::

RESULT 15
US-08-645-193B-18
: Sequence 18, Application US/08645193B
: Patent No. 5962253
: GENERAL INFORMATION:
: APPLICANT: Kupke, Thomas
: APPLICANT: Gotz, Friedrich
: APPLICANT: Kempler, Christoph
: APPLICANT: Jung, Gunther
: TITLE OF INVENTION: Oxidative Decarboxylation of Peptides
: NUMBER OF SEQUENCES: 70
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
: STREET: 1100 New York Avenue, Suite 600
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/645,193B
: FILING DATE: 13-MAY-1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Esmund, Robert W.
: REGISTRATION NUMBER: 32,893
: REFERENCE/DOCKET NUMBER: 0652.1540000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: INFORMATION FOR SEQ ID NO: 18:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 8700 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: both
: MOLECULE TYPE: cDNA
: US-08-645-193B-18

Alignment Scores:
Pred. No.: 0.00454 Length: 8700
Score: 123.00 Matches: 153
Percent Similarity: 32.468 Conservative: 120
Best Local Similarity: 18.198 Mismatches: 258
Query Match: 2.87% Indels: 310
DB: 2 Gaps: 40

US-09-978-385-2 (1-805) x US-08-645-193B-18 (1-8700)
Qy 1 MetSerSerSerTTPleuLeuLeuSerLeuValAlaValAlaIleGlnSerThr 20
      :::::
Db 2662 ATTAGTGAAGCAGCATATATCTCTGCTATTATCTCTAATCATTTTGGTACAAAAC 2621
      :::::
Qy 21 IleGluGluGlnAlaIleThrPheLeuAspLysPheAsnHisGlu----- 35
      :::::
Db 2622 ATTAGGAATTATTCAGCAATTTTATGATTAATATGATTTGAACAACAGTAGTAATTTA 2681
      :::::

```

```

Qy 36 -----AlaGluAspLeu-----PheTyrGlnSerSerLeuAlaSerTrpAsnTyr 50
      :::::
Db 2682 AACCAATGCTCTCAGCATATTAATGATTTGGCTATCCCAAAAAGACAGTTATAGTTT 2741
      :::::
Qy 51 AsnThrAsnIleThrGluGluAsnValGlnAsnMetAsnAlaGlyAspLysTrpSer 70
      :::::
Db 2742 TCTAATACAT----- 2753
      :::::
Qy 71 AlaPheLeuLysGluGlnSerThrLeuAlaGlnMetTyrProLeuGlnGluIleGlnAsn 90
      :::::
Db 2754 GCATTTTAAAGAAAAG----- 2771
      :::::
Qy 91 LeuThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGlu 110
      :::::
Db 2772 -----TATTGCTTGCATTTCAATTAACAGCCATTTGAAATTAACAGAA 2816
      :::::
Qy 111 AspLysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGly 130
      :::::
Db 2817 AACGACGTTAAATAATTTAGAAAAG---AATTAATACAGTTTCAATAATCAATGCG----- 2867
      :::::
Qy 131 LysValCysAsnProAspAsnProGlnGluCysLeuLeuLeuGluProGlyLeuAsnGlu 150
      :::::
Db 2868 ----- 2882
      :::::
Qy 151 IleMetAlaAsnSerLeuAspTyrAsnGluArgLeuTyrPalaIleProGluSerTrpArgSer 170
      :::::
Db 2883 ATA-----TATAGTCGATATATTT----- 2903
      :::::
Qy 171 GluValAlaGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGluMet 190
      :::::
Db 2904 -----GGAATTCATTAATAAGGT---TATGAGATTTGGCGGTAGTCAATATATTA 2954
      :::::
Qy 191 AlaArgAlaAsnHisTyrLysAspTyrGlyAspTyrTrpArgGlyAspTyrGluValAsn 210
      :::::
Db 2955 GGATCTTTAAAGCCCGCGTCAACTTTGGAAGCTTT---ACGGAAATTTCAATTAAG 3011
      :::::
Qy 211 GlyValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGlnAspValGluHisThrPhe 230
      :::::
Db 3012 -----AAAAAATCAATTAACAAAAAGAAATAGTCAT----- 3044
      :::::
Qy 231 GluGluIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsn 250
      :::::
Db 3045 -----CATTCATTAATTAATCAATGAAGAAGAAAGTTAGAA 3080
      :::::
Qy 251 AlaTyrProSerTyrTyrLeuSerProIleGlyCysLeuProAlaHisLeuLeuGlyAspMet 270
      :::::
Db 3081 ATTAGCCAAATTAATGAAGGTCCTTAATCTACAGAAATGTAATAATTTTGAATAT--- 3137
      :::::
Qy 271 TrpGlyArgPheTrpThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsn 290
      :::::
Db 3138 -----AATGAAATATATATCTGTTAAATTTAAATTTACCT-----AAAAGTGAT 3185
      :::::
Qy 291 IleAspValThrAspAlaMetValAspGlnAlaTyrPaspAlaGlnArgIlePheLysGlu 310
      :::::
Db 3186 ATGAGATATAAGACATATATTATGAGCTACATTAACAACACTTATCTATATCTGAA 3245
      :::::
Qy 311 -----AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnIle--- 326
      :::::
Db 3246 AACATGATTCAGAAATGATTTGATCTAATTAATTCATTAATTAATGAGCTTGAGACT 3305
      :::::
Qy 327 -----PheTyrGluAsnSerMetLeuThrAspProGly 337
      :::::
Db 3306 GAATTTACAAATTTTAAAGAGAAATTTGATTTGAAAAACAAATTTATACAACTATA 3365
      :::::
Qy 338 Asn-----ValGlnLysAlaVal 343
      :::::
Db 3366 ACTGAAGAGGATGACTCATTAACCTTTTGTCCAAAGAAATTTATTAATAAATTTAT 3425
      :::::
Qy 344 CysHisProThrAlaThrAspLeuGlyLysGlyAspPheArgIleLeuMetCysThrLys 363
      :::::
Db 3426 TTAATAACAGCTACTTGGAAATTAATTAATTCAGAAATGTTTTCGAA-----ACTGA 3476
      :::::
Qy 364 ValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGlnTyrAspMet 383
      :::::

```

```

Db 3477 AATGGTTAAATAGCTTCGCAACTATT---AGAGAAAAAATGCATATTCGAAAAAGATGA 3533
QY 384 AlAtyRaA----- 386
Db 3534 ATTATTCGCTTTGGAGATATTCGATTCGTAATTAATTATTAATGACAGCATCTCAT 3593
QY 387 -----AlaGlnProPheLeuArgAsnGlyAlaAsn 397
Db 3594 ATACTAAAAAGAACTAAAAAACAATGATGATTCGATTCGATTAAGAAAGCTTATCAAT 3653
QY 398 GlnGlyPheHisGluAlaValGlyGluIleMetSer----- 409
Db 3654 GAATCTAATATGAGAGAAATGTAAGAAATGTTACGCCATTAATATATATATTAATGTTA 3713
QY 410 -----LeuSerAlaAlaThrProLys-----HisLeuLys--- 419
Db 3714 AAAGAACAAATCTTCATATACCTAAATAAATAAGAAATAGACATTCATATATCTTAAAGAT 3773
QY 420 -----SerIleGlyLeuLeuSerProAspPheGlnGlnAspAsnGluThrGluIleAsn 437
Db 3774 TGGTTTCAATTCATTTAATGATTCCTAAACATACCAAGAT-----AAT 3818
QY 438 PheLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrlMetLeuGlu 457
Db 3819 TTTATTCAGATTATCTATACCATTTATACGGAATTAAGTTAATATATTTATTTAAT 3878
QY 458 LysThrArgThrMetValPheLysGlyGluIleProLysAspGlnThrMetLysLysThr 477
Db 3879 AAATTTTTCATCAATAAATTTAAAGAGAT-----GAAATTTTAAATAA----- 3923
QY 478 TrpGluMetLysArgGluIleValGlyValGluProValProHisAspGluThrTyrl 497
Db 3924 TTAAGATTATTAAGAGA-----GATGAAGATTAT 3953
QY 498 CysAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrlr--- 516
Db 3954 -----TCCAATTTATTCCTTCATTAATAAATAATTCGAAA 3986
QY 517 -----ThrArgThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuLysGln 531
Db 3987 GATTATTCCTTAAATAGTAGAATTATATGACTATTCATTAATGATGTTCTCTGAA 4046
QY 532 AlaAlaLysHisGlnGlyProLeuHisLysCysAspLysSerAsnSerThrGluAlaGly 551
Db 4047 GATATATAGATATGCTGCACACGTAATTAAGATATGAGAAATTTTATGATGAT 4106
QY 552 GlnLysLeuPheAsnMetLeuArgGlyLysSerGluProThrThrLeuAlaLeuGlu 571
Db 4107 AGTCTATTA-----TCATTAATATATATACATCAAG--TTCAAAATTCGAAAAGAA 4157
QY 572 AsnValValGlyAlaLysAsnMetAsnValArgProLeuAsnThrPheGlnProLeu 591
Db 4158 TTTATTCGTTCCT-----ATATCAATATGATTTTATTTGATTAATTTAGAA----- 4202
QY 592 PheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGlyTrpSerThrAspTrpSer 611
Db 4202 ----- 4202
QY 612 ProTyrAlaAspGlnSerIleLysValArgIleSerLeuLysSerAlaLeuLysPlys 631
Db 4203 -----ATTAAATTAAGTAGAAGAAAGAAATTTTAATTAATTAATCCG-----GAAAGAT 4250
QY 632 AlaTyrGlnTrpAsnAspAsnGlnMetTyrLeuPheArgSerSerValAlaTyrAlaMet 651
Db 4251 TTTATTCGATGATATAC-----ATA 4271
QY 652 ArgGlnTyrPheLeuLysValLysAsnGlnMetIleLeuPheGlyGlnGluLysPvalArg 671
Db 4272 AGAGAAATAT----- 4280
QY 672 ValAlaAsnLeuLysProArgIleSerPheAsnPhePheValThrAlaProLysAsnVal 691

```

```

Db 4281 ----AAAAATTATAGCTAAACTT-----ACCAATCTAAAAATGAC 4319
QY 692 SerAspIleIlePro----- 696
Db 4320 TATGAATTTTAAAAAAGAAATTTCCGAATCTTCATGATTTCTTTAATAAATTAAT 4379
QY 697 -----ArgThrGluValGlnLysAlaIleArgMetSerArgSerArgIle 711
Db 4380 ATTTTGAATAATCTTAAAAAGACACTACAAAAAAGCTTATATCTTACAGTTCAGATA 4439
QY 712 AsnAspAlaPhe-----ArgLeuAsnAspAsnSerLeuGluPheGlyIleGln 728
Db 4440 ATTCGACGTTTATACATGCGTTGTAAT-----AGAAATTCGATTAAT 4487
QY 729 Pro 729
Db 4488 CCT 4490

RESULT 16
US-09-280-116-114/C
; Sequence 114, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 114
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: angiotensin-converting enzyme
; NAME/KEY: misc.feature
; LOCATION: (1)-(789)
; OTHER INFORMATION: n = a, t, c or g
US-09-280-116-114

Alignment Scores:
Pred. No.: 0.000133 Length: 789
Score: 121.00 Matches: 31
Percent Similarity: 47.78% Conservative: 12
Best Local Similarity: 34.44% Mismatches: 29
Query Match: 2.82% Indels: 18
Gaps: 3

US-09-978-385-2 (1-805) x US-09-280-116-114 (1-789)
QY 186 LeuLysAsnGlnMetAlaArgAla-----AsnHisTyrGluAsp 198
Db 384 CTAAGGCGTAGAGATGCTGAGCTCTGAGAGTCTCTCTCTGCTAGGTACAAAGAC 325
QY 199 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlnValAspGlyTyrAspTyrSer 218
Db 324 ATGGGGCCCTTGCGCACTCCAAATATGATGAGTCCGATCCCTGGAG----- 280
QY 219 ArgGlyGlnLeuIleGluAspValGlnHisThrPheGlnGluIleLysProLeuTyrGlu 238
Db 279 -----CAAGACCTTGAGCGCGGTATTTCCAGAGAGCTGGCGGCACTCTACCTG 235
QY 239 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyr---ProSerTyrIleSer 257
Db 234 AACCCGACACCTAGCTGCGGAGGCGCCCTCCACCGCCACATATGGGCGGAGCTCATGAC 175
QY 258 ProIleGlyCysLeuProAlaHisLeuLeu 267
Db 174 CTGAGGGGGCCCATCCCTGCCCACTCTCTG 145

RESULT 17

```

US-09-453-702B-1/C
 ; Sequence 1, Application US/09453702B
 ; Patent No. 6365723
 ; GENERAL INFORMATION:
 ; APPLICANT: Blatner, Frederick R.
 ; Burland, Valerie
 ; Perna, Nicole T.
 ; Plunkett, Guy
 ; Welch, Rod
 ; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
 ; NUMBER OF SEQUENCES: 265
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Quarles & Brady
 ; STREET: 1 South Plunkney Street
 ; CITY: Madison
 ; STATE: WI
 ; COUNTRY: US
 ; ZIP: 53701-2113
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 8.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/453,702B
 ; FILING DATE: 03-Dec-1999
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/110,955
 ; FILING DATE: 04-DEC-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seay, Nicholas J.
 ; REGISTRATION NUMBER: 27386
 ; REFERENCE/DOCKET NUMBER: 960296,95017
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (608) 251-5000
 ; TELEFAX: (608) 251-9166
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6506
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 ; US-09-453-702B-1
 ;
 ; Alignment scores:
 ; Pred. No.: 0.00548 Length: 6506
 ; Score: 120.50 Matches: 186
 ; Percent Similarity: 32.57% Conservative: 128
 ; Best Local Similarity: 19.29% Mismatches: 328
 ; Query Match: 2.81% Indels: 323
 ; Gaps: 51
 ;
 ; US-09-978-385-2 (1-805) x US-09-453-702B-1 (1-6506)
 ;
 ; QY 5 SertripleuLeuSerleuValAlaValThraAlaInserThrIleGluInGln 24
 ; Db 3778 ACATTATTCGACACCGCCAGTAAGTCGTGTAGCAGAGTGTGCGCAACAATAA 3719
 ; QY 25 AlalysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuTherYrGlnSerSer 44
 ; Db 3718 ATAGAAAAATGCCAAACAATTGGCTCCACAAATAAGAAAGTTTATTATTAT 3668
 ; QY 45 LeuAlaSerTrpAsnTrpAsnThrAsnIleThrGluGluAsnValGlnAsnMetAsnAsn 64
 ; Db 3667 -----CTAAATGTCTGATATCCCTCCCAACAATCCACAAAT 3629
 ; QY 65 Alagly---AspLysTrpSerAlaPheLeuLysGluGlnSerThreU----- 79
 ; Db 3628 GCAGGAAAAACAATAATTAACTTGATGCAAAACCGGATAAAGTTACTTGGCGTCA 3569

QY 80 AlagInMetYrProleu-----GlnGluIleGlnAsnLeuThrValLysLeuGlnLeu 97
 ; Db 3568 AGCGAATACCTCCGCTAGATAAAAAAACCTTCTCATTAATATACAAAAGAAAAT 3509
 ; QY 98 GlAlaLeuGlnGlnAsnGlnSerSerValLeuSerGluAspLysSerLysArgLeuAsn 117
 ; Db 3508 AACGCATCTCAATAAACAACGAAACAGCC-----AAC 3476
 ; QY 118 ThrIleLeuAsnThrMetSerThrIleYrSerThrGlnLysValCysAsnProAspAsn 137
 ; Db 3475 TGGATA-----ACGGTTACACTATC-----AAGCAACAAGAGTAAAGTT 3434
 ; QY 138 ProGlnGluCysLeuLeuLeuPro---GlyLeuAsnGluIle---MetAlaAsnSer 155
 ; Db 3433 AATAATGAAGAATATCTTACTGCTCCCTTCGCTAAGCAACATTAATAAATAATAC 3374
 ; QY 156 LeuAspTrpAsnGluArgLeuTrpAlaTrpGlnSerTrpArgSerGluValGlyGln 175
 ; Db 3373 CACGCATCGAATATGAACTAAGCTAGTATGATGAGGGAATATACATCCATATGATA 3314
 ; QY 176 LeuArgProLeuTrpGluGluTrpValValLeuLysAsnGln-----MetAlaArgAla 193
 ; Db 3313 ATCGCCGCGAGATTAACAGGAAAAAGTCATGTTAAGGACACAGATGGTGGCTGAT 3254
 ; QY 194 AsnHisTrpGluAspTrpGlyAspTrpTrpArgGlyAspTrpGlu-----Val 209
 ; Db 3253 ATTTTGTGTATTCCTTCCGCGGTATGCTGAAGAACTTTTGAATGCACTTATGATC 3194
 ; QY 210 AsnGlnValAspGlyTrpAspTrpSerArgGlyGlnLeuIleGluAspValGluHisThr 229
 ; Db 3193 GCGCGCATGAAGAAGGAAAAAGTTCC-----GATATCAT 3158
 ; QY 230 PheGluGluIleLysProleu-----TrpGlnHisLeuHisAlaTrpVal----- 244
 ; Db 3157 TTGATATTAATAACACCCCTTCCAGGAACATACAA---CTGATTTTATGTAATAATAC 3101
 ; QY 245 -----ArgAlaLysLeuMetAsnAlaLysProSerTrpIleSerProle---GlyCys 261
 ; Db 3100 CAGTGGCGTGAAGAAACAAGATATACAGATCCCAAG-----TCACCATGCAAAACCATGT 3047
 ; QY 262 LeuProAlaHisLeu-----LeuGlyAspMetTrpGlyArgPheTrpAsn--- 277
 ; Db 3046 CTGCCTAAGTGTACTATCAACCAACCTGGGGGTAAGACGGGTAAACCTCATACAGAAGT 2987
 ; QY 278 -----LeuTrpSerLeuThrValProPheGlyGln----- 287
 ; Db 2986 AATTGTATTATTCGATTAAGCGGTTACGTGAGCAATATACATGCGATATCACTGAA 2927
 ; QY 288 ---LysProAsnIleAspValThrAspAlaMetValAsp----- 299
 ; Db 2926 CACGCACTAATTTAACAAGTACCAAGCCATATATCAATGAACTGAAAGAGTTACGTT 2867
 ; QY 300 -----GlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSer 317
 ; Db 2866 CCCCCGGAAGAGTTGGAT-----CGCGCATGATGCTTTTATACCTCC 2822
 ; QY 318 ValGlyLeu----- 320
 ; Db 2821 TATATTTAAGTCACTGCTGCTTATGACAGCAATATATACATATACAGCCAGTTAT 2762
 ; QY 320 ----- 320
 ; Db 2761 GCGCATTTAACAAGTGAATTAATTTATAGTTGACACTTACACTGATCCAGCTAT 2702
 ; QY 321 ---ProAsnMetThrGlnGlyPheTrpGlu---AsnSerMetLeuThrAspProGly 337
 ; Db 2701 AGCAAAACCGATGATATGAAGACACATGCGCAAAATTAATACCTTACCTGGAACATGCG 2642
 ; QY 338 -----AspValGlnHisAlaValCysHis 345
 ; Db 2641 TGTGCGAGATATTAAACACAGTTCAGATAGGAAATAATACATCCCTCATCTATATTC 2582
 ; QY 346 ProThrAlaTrpAspLeuGlnLysGlyAspPheArgIleLeuMetCysThrLysValThr 365

Db 2581 GACCTCTGAGATTAGCGGAATACGTTATTCGCGATATGCAATGCTTACCTGATTCCT 2522
 QY 366 MetLysAspSphE-----LeuThrAlaHis 373
 Db 2521 ATGCGAGGCTTTACACCTTTAGTACAGGCTTGCCGCAAGTAATGCCCTTGTAACGGTT 2462
 QY 374 HisGluMetGlyHisIleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuArg 393
 Db 2461 TCGCAAAATGGCTACATCATTTACCAAAAAGAGTTCCCGGAGCATTTACTATCGCA 2402
 QY 394 AsnGlyAlaAsnGlyGly-----PheHisGluAlaValGly 405
 Db 2401 GATTTCGACACTCTCGCGACCGGCTGATCTGATGTCAGCTAAAGACAGATGCGC 2342
 QY 406 GluIleMetSerLeu-----SerAlaIleThrProLysHisLeuSerIleGly 422
 Db 2341 AGTGTGCTTCCTTTTGGTGCCCTACTCTCCGTGCTTAACATGCTGCAACCT---GGT 2285
 QY 423 LeuLeuSerProAspSphE-----GlnGluAsp 431
 Db 2284 ATTCAATATTTCGATTTTATGCGCGAGCAGATATATGCTGTAAGAATCAGAG 2225
 QY 432 AsnGluThrGluIleAsnPheLeu-----LeuLysGlnAlaLeuThrIleValGly--- 448
 Db 2224 GACTTTCGAGCGCAAAATATATATATATGAGCTCAACATCTTTGACACTTATGCGCGT 2165
 QY 449 ---ThrLeu-ProPheThrTyrMetLeuGlyLysTyrArgTyrMetValPheLysGly 467
 Db 2164 ACGATTATTCGATTAACATAATGCTATAACTCTGG--- 2128
 QY 467 uIleProLysAspGlnIlePheLysLysTyrPheGluMetLysArgGluIleValGly 487
 Db 2127 -----GAAATGGTGGGAAT 2114
 QY 487 IValGluProValPro-----HisAspGluThrTyrCysAspProAlaLase 502
 Db 2113 ACACCGCTGGGTGCCATCTCTTTGAGCGTACAGATCAGATGAGTAGTAATGTA 2055
 QY 502 rLeuPheHisValSerAsnAspTyr-----SerPheIleArgTyrTyrThrArgHile 520
 Db 2054 TATTAACACAGCAAGGTACAGCTATACAGTCCCTATATATAATATCTAGTAAACGGC 1995
 QY 520 uTyrGlnPheGln-----PheGlnG 527
 Db 1994 CACTCGATTAGCGTGGCCGCTGCTTTCGAGGATTACAGCATTCAGCGCA 1935
 QY 527 uAlaLeuGlyGlnAlaAlaValHisGluGlyProLeuHisLysCysAspIleSerAsn 547
 Db 1934 CCATCTTTACGAAATGATTAATAATATAT-----CATCAGAGTACATGACGATTT 1884
 QY 547 rThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGlnProTyrPth 567
 Db 1883 TTATGCTATTTGGCAGAAAA-----AATAGCCTTTCGCCAATTCATGCCACT--- 1835
 QY 567 rLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValAlaProLeu---LeuAs 586
 Db 1834 -----TTATCAATATATCTGGGA-----AATGATATCATTAGTGGCTTGCGCGAA 1788
 QY 586 nTyrPheGluProLeuPheThrTyrPheLysAspGlnAsnLysAsnSerPhe---ValG 605
 Db 1787 TTAC-----TGCGGCGCAAGTGAATGCTAAAGATTACCAATTACG 1746
 QY 605 yTyrPserThrAspTyrPserProTyrAlaAspGlnSerIleLysValArgIleSerLeu 625
 Db 1745 TTATTTCAATACCTGGCAA-----CGCATCAAT----- 1718
 QY 625 sSerAlaLeuGlyAspLysAlaTyrGluTyrPAsnAspAsnGluMetLysLeuPheArgSe 645
 Db 1717 -----TATCTTCTCTGCG 1704
 QY 645 rSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetIleLeuPhe 665

Db 1703 CAGCAATCCTAT----- 1691
 QY 665 eGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePhe 685
 Db 1690 -GATGAAMCGAC-----AAAGAAGAGAGCGCTTTTAACTGTTAT 1650
 QY 685 lTrhAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAlaIleAr 705
 Db 1649 CTCATTCTCTTCTACTGGGGGATATATTCGCCAAACAGCTCACCAAAATTAATCTATC 1590
 QY 705 gMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPhe 725
 Db 1589 GAATTCACCTCATTTTCCAAAGATGCTATCTTCCTCAACAATACT----- 1544
 QY 725 uGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerIleTyrPheIleVa 745
 Db 1543 -GGATTACTGCGATACCGCTGACATGATGACG-----TTAAA 1506
 QY 745 lPheGlyValValMetGlyValIleValValGlyIleValIleLeuIlePheThrGlyI 765
 Db 1505 TTATGTAATATAT-----GT 1491
 QY 765 eArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsn-----Pr 780
 Db 1490 TAATCAGCAACACAAATAATATATGATCATCGCTTGACGAATTTAAGCTGAGACTGC 1431
 QY 780 oTyrAlaSerIleAspIleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAs 800
 Db 1430 CATGCGCAACATATATGCGCATATGACCTTTTAAACCGCTGCAAAAGTGGGAAG 1371
 QY 800 pValGlnThr 803
 Db 1370 TATATAGAT 1361
 RESULT 18
 US-08-961-527-45/c
 ; Sequence 45, Application US/08961527
 ; Patent No. 6420135
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 391
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/961,527
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PB340P1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 45:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1184 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double

TOPOLOGY: linear
 Alignment Scores:
 Pred. No.: 0.0142
 Score: 120.50
 Percent Similarity: 31.28%
 Best Local Similarity: 17.98%
 Query Match: 2.81%
 Gaps: 45
 US-09-978-385-2 (1-805) x US-08-961-527-45 (1-11384)
 QY 23 GluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPheGln 42
 DB 11096 GAAGCATTAATAAAACCTTTATG-----ACGGGGAATAATTTTATGTCGAA 11052
 QY 43 SerSerLeuLaser----- 47
 DB 11051 CATTATCTAGGAGCACATAGGAGAACTAAATGAGAGCATGCTATACCTTCCTGCTT 10992
 QY 48 TrpAsnTyrAsn-----ThrAsnIleThrGluGlu 57
 DB 10991 TGGGACACCTAATGCTCAGGCTGTCTGCTGCTGATTTACCACTGAGTTGAAAT 10932
 QY 58 AsnValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSer 77
 DB 10931 CAGATTCCAAATGCTAGAAATGATTTTGGGCTGGGAAAGCTTT-----ACC 10884
 QY 78 ThrLeuAlaGln-----MethylProLeuGlnIleGlnAsnLeuThrValLys 94
 DB 10883 AATATGCTCCAGAGAGGAGGATTTTAC----- 10857
 QY 95 LeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLys 114
 DB 10856 ---AATATCATGCTCAGCAGTCAAAATGCTCATCAACGATGAG- 10815
 QY 115 ArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsn 134
 DB 10814 -----ATTGACCTTTTGTCTGACGATGAGGCT- 10785
 QY 135 ProAspAsnProGlnGluLysLeuLeuGluProGlyLeuAsnGluIleMetAlaAsn 154
 DB 10784 -----CGTCCAGGAGACAGGGGCAATCGTAACA- 10758
 QY 155 SerLeuAspTyrAsnGluArgLeuTrp-----AlaTrpGlnSerTrpArgSerGlu 171
 DB 10757 -----GAGCTTCCTGAGAGAAATGGAAGATGAGACTTTGGCTGGCAGCAAGAAACGT 10704
 QY 172 ValGlyLysGlnLeuArgProLeuTyrGluIleTyrValValLeuLysAsnGluMetAla 191
 DB 10703 TGGGGCTTTGGAAGAGCGCTCTCTC----- 10680
 QY 192 ArgAlaAsnHisTyrGlnAspTyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGly 211
 DB 10679 -----AATATTATGAGAGTTCACCGCTGATCATGGAAA- 10641
 QY 212 ValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGlnHisThrPheGlu 231
 DB 10640 TCTGATGGCAGCTCTTATAGTTTGGCCAGCTCAAGATGAAATCATCTCTATCTCGTT 10581
 QY 232 GluIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAla 251
 DB 10580 GAAATGAAC-----TATATCTCATTTGAG----- 10557
 QY 252 TyrProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuGluLysAspMetTrp 271
 DB 10556 -----TTTATGCC-----TTGATGTCCCATCTTTGGGCTTGAGTTGG 10518
 QY 272 GlyArgPheTrpThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysPro----- 289
 DB 10517 GGGATCAGCTTATGGGTTACTCTGCTTTAGCATGCTTATGGCCGACACAGAGGATTT 10458

QY 290 -----AsnIleAspValThrAspAlaMetValAsp 299
 DB 10457 CAGATTTCGAGAGAGTGTCTATACCATTAATATTGGGCTT-----ATTGTGAC 10407
 QY 300 GlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGly 319
 DB 10406 ----- 10392
 QY 320 LeuProAsn-----MethylGlnGlyPheTrpGluAsnSerMetLeuThrAspProGly 337
 DB 10391 TTACCATCAACAGATGATGCTTGA-CCCTATTTATGATGGACACGACACTTTGAATCCAA 10333
 QY 338 AsnValGlnLysAlaValLysHisProThrAlaTrp-----AspLeuGly 352
 DB 10332 GACCATTAATTAAGCT-----CATTAACATGCTGGGCTGCCCTTAATTTTGACCTTGA 10279
 QY 353 LysGlyAspPheArg-----IleLeuMetCysThrLysValThrMetAspPhe--- 369
 DB 10278 AAAAATGAAGTCAGTCGCTCTTAATTTCTTCATTAAACATGATGATGATGATGAT 10219
 QY 370 -----LeuThrAlaHisGlnMetGlyHisIleGlnTyrAspMetAla 384
 DB 10218 TTGATGATGATTCGCTGCTGATGCTTTAGCAACATGCTCTATTTGACTATGATGATGCT 10159
 QY 385 TyrAlaAlaGlnProPheLeu---LeuArgAsnGlyAlaAsn-----GluGlyPhe 400
 DB 10158 -----CCATGACACCTTAATTAAGATGGCGGAATCTCACTAATGAAGTTAT 10111
 QY 401 His-----GlnAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 418
 DB 10110 TATTTCTTCAGGCTTGAATGAGGTTTATAGTTAGAA----- 10072
 QY 419 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 438
 DB 10071 -----TATCCAGATGTCATGATGATTCGCAAGAAAGTCTGCTGCGC 10030
 QY 439 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheTrpTyrMetLeuGlnLys 458
 DB 10029 ATCAAGATTACGGGATGAGATGAAGAGATGAGTGTGCTCAGGATTTGACTAC-----AAA 9979
 QY 459 TrpArg-----TrpMet-----ValPheLysGlyLysIlePro----- 469
 DB 9978 TGGACATGGCTGGATGATGATATTCCTCGCTTCTCAAGAAAGATCCGATGTATCGT 9919
 QY 470 -----LysAspGlnTrpMet 474
 DB 9918 AATATGACTTTAACCTGCTGACTTTCAGCTTATGATGTTTTCAGAGAAATATATCTC 9859
 QY 475 LysLysTrpTrpGluMetLysArgGluIleValGlyValGluIleProValProHisAsp 494
 DB 9858 TIG-----CCATTCCTGCGACGAT 9841
 QY 495 GluThrTrpTyrCysAspProAlaSerLeuPheHisValSer-----AsnAsp 509
 DB 9840 GAAGTGGTTCATGGCAGAGAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 9781
 QY 510 TyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnGluAlaLeu 529
 DB 9780 TTGCGAGGCTTG-----CGCAATCTTATACGTAACAA-----ATT 9745
 QY 530 CysGlnAlaAlaLysHis----- 535
 DB 9744 TGTCACTCTGTAGAAATTTCTCTTCAATGGTACCAATACGCTCAATTTCTTGAATGAG 9685
 QY 536 -----GluGlyProLeu----- 539
 DB 9684 AATCTGAAGACAGTTGATGATGCTTAACTAGAACACCAATGAATGCTAAGATGAAG 9625
 QY 540 -----HisLysGly----- 542
 DB 9624 TATTTGCTTCTCAGCTTAACCACTTTTACAAGATCATGCTGCTGTGGGAAATGAT 9565
 QY 543 -----AspIleSerAsnSerThrGluAlaGlnLysLeuPheAsn 556

Query Match:	2.80%	Indels:	298
DB:	4	Gaps:	42
US-09-978-385-2	(1-805)	x	US-08-961-527-115 (1-11303)

OY	40	PheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThr--AsnIleThrGlnLysAsn	58
Db	2860	TTTTTTCAATGGACTATTATTGTCTCAGTAGTAACACAGATAACAACATTGCATTCAGATA	2801
OY	59	ValGlnMetLeuMetAsnAlaGlyAspLysTrpSerAlaPheLeu-----LysGlu	75
Db	2800	ATGAAATAAACCATCAAAAAATG-----GTCCGCTTCGTTTTATCTCATCAATAAAAAGAG	2747
OY	76	GlnSerThrLeu-----AlaGlnMetTyrProLeuGlnGluIleGlnAsnLeu	91
Db	2746	GAATAACAATVTGAATAATTTTAAAGGAGAAGTTTAT-----CGAATCTCTGTATTA	2696
OY	92	ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAsp	111
Db	2695	ACAACAAGCTAGTAAAGTTAAGTAGAATCTCTCAAACTGGAAGT-----TTTGAGCAT	2645
OY	112	LysSerLysArgLeu-----	116
Db	2644	AGAACGACACAACTTATCTTAATTAAGAATTTTGGCCAATTTGCTTAGATTATTAAGAG	2585
OY	117	--AsnThrIleLeuAsnThrMetSerThrIle-----TyrSerThrGlyLys	131
Db	2584	ACATCAAACTACTAGATATTATATGACGAGCATTTTTCATCTCCACTTAAATTAAGAGAA	2525
OY	132	ValCysAsnProAspAsnProGlnGlnCysLeuLeuLeuGlnProGlyLeuAsnGluIle	151
Db	2524	TTT--AACCCCAAAAT-----TTATTATTA-----GAATTA	2495
OY	152	MetaIasnSerLeuAspTyrAsnGlnLysArgLeuTrpAlaTrpGluSerTrpArgSerGlu	171
Db	2494	AAAGAAATTTTCCCGTAATGTGTATGCCCTGTATTTTGGTGAATCTATTGAACGTTA	2435
OY	172	ValGlnLysGlnLeuArgProLeu-----TyrGlnGluIleTyr	183
Db	2434	AAAAGA---ACACCTCGGACTCTGGATTAAGGACAGATGAGCAATCTCGTGAAGATGGA	2378
OY	184	ValValLeuLysAsnGlnMetAlaValGalaIsnHisTyrGlnAspTyrGlyAspArgTyr	203
Db	2377	ATTATTAGCCGAATGTGTATAGCTTATTGGAT-----GATTCGAAGATTATTATGG	2324
OY	204	----ArgGlyAspTyr-----GluValAsnGlyValAsp-----	213
Db	2323	GATGAACAATCTGGTATTATATGAGACAGAAATCAATGATGACCTGTATTCTTGGCAT	2264
OY	214	GlyTyrAspArgTyrSerArgGlyGlnLeuIleGluAspValGlnHisThrPheGluGluIle	233
Db	2263	GGGGATGATTAAT--AGAGGA---GCAATCAAGAGATTTTACCAAT-----	2225
OY	234	LysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrPro	253
Db	2224	-----	2207
OY	254	SerTyrIleSerProIleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArg	273
Db	2206	-----TTGTTGCCCAAAATATGCTTATGGCAAAATTTGGTGGAGTGA	2168
OY	274	PheTrpThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsnIleAspVal	293
Db	2167	TATTGG-----CCTTATAGCTGGATGAATCTTGATTTA	2132
OY	294	ThrAspAlaMetValAspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLys	313
Db	2131	ATGACACGATTT-----GAAACAGAGAAA	2108
OY	314	PhePheValSerValGlyLeuProAsnMet-----	323
Db	2107	ATTGCATTAATCTATCGGTGTGTGATTAGATTGGCATATTACTGACATTCACACTCGC	2048

Oy	324	-----	ThrGlnGlyPheProGluAsnSerMetLeuThrAspProGlyAsn	338
Db	2047	TTTGGAAAGTGCCTGCACAGCATCTAGTTGGAAATAGAACTTAATACCAATCCAAACACAG		19886
Oy	339	ValGlnIysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle		358
Db	1987	TTATATGCACACACTT---CATGATATGAAAGCTTAAACTCTCC-----		19494
Oy	339	LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGlnMetGlyHis		378
Db	1948	-----	TTAAATGTCACCTCTGATGGGATA	19222
Oy	379	IleGlnIyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu		398
Db	1921	CGGGCTTATGACGAAGCTTAT-----		19011
Oy	399	GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu		418
Db	1900	-----	CCCTAAATCCGA	18882
Oy	419	LysSerIleGlyLeu-----	LeuSerPro	426
Db	1888	AAACGGTTGGGGTTAAATGTAGAACTAGAAAGACCTGCTATTTTTCATTTTATATCC		18222
Oy	427	AspPheGlnGluAspAsnGlnThrGluIleAsnPheLeuLeuLysGlnAlaLeuThrIle		446
Db	1828	TCTTTTGGGAGAGCTTACTTAAAGCATGTTCAT-----		17966
Oy	447	ValGlyThrLeuProPheThrTyrMetLeuGlnLysTrpArgTrpMetValPheLysGly		466
Db	1795	-----	TATGAACATAGAAAG-	17811
Oy	467	GluIleProLysAspGlnTrpMetLysLysTrpTrp---GluMetLysArgGluIleVal		485
Db	1780	-----	CAGGAGATAGATTTTGGTGGATGACTGGCAGACAGACAGACAA	17366
Oy	486	GlyValValGluProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHis		505
Db	1735	GGATGCTAATCCACTT-----		1718
Oy	506	ValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPhe		525
Db	1717	-----	TGGCTTTTAAACCATATACATAT	16944
Oy	526	GlnGluAlaLeuCysGlnAla-----	AlaLysHisGluIlePro---	538
Db	1693	CAGATATGTTGTAAAAATGCAGAAAGGTGGTGTATTTTCAAGATATCGAGAGCTCTGGT		16344
Oy	539	LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeu-----PheAsn		556
Db	1633	AGTCACCGCATCCCTGTGGTGTTCAT-----	GGGATATCATATTATATAGTTGGAT	15833
Oy	557	MetLeuArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGlnAsnValValGlyAla		576
Db	1582	TCCTTAAGATT-----	CAACCTATTTTACGCGACAGCATATATATCCGT-	15355
Oy	577	LysAsnMetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLys		596
Db	1534	-----	TATAGTTGG-----	15266
Oy	597	AspGlnAsnLysAsnSerPheValGlyTyrSerThrAspTrpSerProTyrAla-----		614
Db	1525	-----	TGGAGTCATGATATCGGAGCAATATGTTGGGG	14933
Oy	615	-----AspGlnSerIleLysValArg-----	IleSerLeuLysSer	626
Db	1492	GATTATACGACAGACGACAACTAGATGCGTACAGTTTGCTGTTTATAGTCCGATTACT		14333
Oy	627	AlaLeuGlyAspLysAlaTyrGlnTrpAsnAspAsnGluMetTyrLeuPheArgSerSer		646
Db	1432	CGATTACATAGTCTTAGAAGTCTCTTTATATAGTAAAGAACCTGGTTTTTTTCAGAAACA		13733
Oy	647	ValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetIle	663	

Fri Mar 14 10:00:44 2003

us-09-978-385-2.p2n.rn

Page 24

```

Db 1372 ACATCTAAGATTATGAGAGAAATAC--CTTCGTTTGAGACATCAGATGATT 1325
      ::      |||:::|||||      |||:::|||||

```

Search completed: March 7, 2003, 09:50:13
Job time : 166 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 7, 2003, 09:09:04 ; Search time 163 Seconds
(without alignments)
3298.623 Million cell updates/sec

Title: US-09-978-385-2

Sequence: 1 MSSSSWLLSLVAATAAOST.....ISKGNNGPQNTDVCYNSF 805

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 478924 segs, 333959956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ .p2n.model -DEV-rlp
-Q=/cgn2_1/USPTO.spool/US09978385/runat_28022003.104718.1445/app_query.fasta_1.967
-DB-published.Applications.NA -QEMT-fastlap -SUFFIX-p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS-bits -START=1 -END=-1 -MATRIX-blosum62
-TRANS-human40.ccd -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=20 -MODE-LOCAL -OUTFMT-pto -NORM-ext -HEAPSIZE=500 -MUTLEN=0
-MAXLEN=2000000000 -USER=US09978385 -GCCN_1_1.80 -runat_28022003.104718.1445
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications.NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/ECT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/ECTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query length	ID	Description
1	4291	100.0	3334	9	US-09-978-385-1
2	4287	99.9	3325	9	US-10-114-893-85
3	4142	96.5	3732	9	US-10-028-072-71
4	4142	96.5	3732	9	US-10-121-049-71

Result No.	Score	Match	Query length	ID	Description
5	4142	96.5	3732	9	US-10-123-904-71
6	4142	96.5	3732	9	US-10-140-470-71
7	4142	96.5	3732	9	US-10-175-746-71
8	4142	96.5	3732	9	US-10-176-921-71
9	4142	96.5	3732	9	US-10-176-921-71
10	4142	96.5	3732	9	US-10-137-865-71
11	4142	96.5	3732	9	US-10-140-474-71
12	4142	96.5	3732	9	US-10-142-431-71
13	4142	96.5	3732	9	US-10-143-114-71
14	4142	96.5	3732	9	US-10-140-002-71
15	4061	94.6	2920	9	US-09-969-384-2
16	4013	93.5	2911	9	US-09-969-384-12
17	3579	83.4	2638	9	US-09-978-385-5
18	3561	83.0	2638	9	US-09-978-385-8
19	3509	81.8	2415	9	US-09-978-385-3
20	2504	67.7	2415	9	US-09-978-385-7
21	1337	31.2	4022	9	US-09-964-899-18
22	721	16.8	2082	10	US-09-864-761-10784
23	376.5	8.8	456	10	US-09-992-598-386
24	376	8.8	1346	9	US-09-988-283A-386
25	376	8.8	1346	9	US-09-983-735-386
26	376	8.8	1346	9	US-09-990-444-386
27	376	8.8	1346	9	US-09-989-730-386
28	376	8.8	1346	9	US-09-990-436-386
29	376	8.8	1346	9	US-09-991-181-386
30	376	8.8	1346	9	US-09-993-687-386
31	376	8.8	1346	9	US-09-989-734-386
32	376	8.8	1346	9	US-10-028-072-481
33	376	8.8	1346	9	US-09-997-653-386
34	376	8.8	1346	9	US-09-993-667-386
35	376	8.8	1346	9	US-10-121-049-481
36	376	8.8	1346	9	US-10-123-904-481
37	376	8.8	1346	9	US-10-140-470-481
38	376	8.8	1346	9	US-09-990-438-386
39	376	8.8	1346	9	US-09-990-562-386
40	376	8.8	1346	9	US-09-997-428-386
41	376	8.8	1346	9	US-09-997-428-386
42	376	8.8	1346	9	US-09-997-428-386
43	376	8.8	1346	9	US-10-175-746-481
44	376	8.8	1346	9	US-10-176-921-481
45	376	8.8	1346	9	US-10-176-921-481

ALIGNMENTS

RESULT 1
US-09-978-385-1
; Sequence 1, Application US/09978385
; Patent No. US2002017211A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Petrie, Charles
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME
; FILE REFERENCE: 99-24C1
; CURRENT APPLICATION NUMBER: US/09/978,385
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/135,952
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 60/151,181
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/563,516
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3334
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)...(2449)

US-09-978-385-1

Alignment Scores:

Pred. No.: 0 Length: 3334
 Score: 4291.00 Matches: 805
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-978-385-2 (1-805) x US-09-978-385-1 (1-3334)

QY 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
 DB 35 AATGCAAGCTCTCCCTGCGCTCCTTCACACCTGTGGTGTGAACGTGGCTCAGTCCAC 94
 QY 21 IleGluGluAlaAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
 DB 95 ATTGAGAAACAGGCCAAGACATTTTGGACAAAGTTTAAACACAGAACCCGAAACCTGCTTC 154
 QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsnValGln 60
 DB 155 TATCAAAAGTTCACCTGCTCTTGGATATTAACACCAATATTACTGAAAGAGATGCCAA 214
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
 DB 215 AACATGAAATAAAGCTGGGGACAAATGGTCTGCTTTTAAAGAACAGTCCACACTGCC 274
 QY 81 GlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
 DB 275 CAAAGTATACACTACAAACAAATTCAGAAATTCACAGTCAAGCTTCACCTGACGGCTCTT 334
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
 DB 335 CAGCAAAAGGGTCTTCACCTGCTCTCAGAGACAGACAGAAAGCGTTGACACCAATTCTA 394
 QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
 DB 395 AATACATATGACACATCATCTACAGTACTGGAAGTTTGAACCCAGATTAATCCACAA 454
 QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
 DB 455 TCCCTATTAATCTGAACACAGGTTTGAATGAATGAATGCAAAACAGTTTGACATACATAG 514
 QY 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyr 180
 DB 515 AAGGCTGGGCTGGGAAAGCTGGAGATCTGAGTCCGCAACAGCTGAGGCCATTAAT 574
 QY 181 GluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200
 DB 575 GAAGAGTATGTGCTTGAATAATGAGATGGCAAGCAAAATCAATTATGAGACTATGG 634
 QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
 DB 635 GATTTATGGAGAGGAGACTATGAATGAATGAATGGGTAGATGCTATGACTACAGCCGCGC 694
 QY 221 GlnLeuIleGluAspValGluIleThrPheGlnGluIleLysProLeuTyrGluHisLeu 240
 DB 695 CAGTTGATTGAAGATGTGAACATACCTTTGAAGATTTAAACCTTTATATGAACTATT 754
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
 DB 755 CATGCTATGTGAGGGCAAGTTGATGATGCTATCCTTCTATATGAGTCCAAATTGA 814
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280
 DB 815 TCCCTCCCTGCTCATTTGCTGGTATATGTGGGTAGATTTTGGACAAATCTGACTCT 874
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
 DB 875 TTGACAGTTCCCTTTGGACAGAAACCAACATAGATGTACTGATGCAATGTGTGACCG 934
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320

DB 935 GCCTGGCATGCACAGAAATATTCAGAGAGCCGAGAACTTCTTGATCTGTGGCTT 994
 QY 321 ProAsnMetTrpGlnGlyPheTrpGluAsnSerMetLeuThrAspProGluAsnValGln 340
 DB 995 CCTATATATGACTCAAGATATCTGGGAAATTTCCATGCTAAGCAGACCCAGGAAATGTTCCAG 1054
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
 DB 1055 AAAGCAGCTGCTCCATCCACAGCTTGGGACCTGGGAGAGGCCACTTCAGAGATCCTTATG 1114
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHisIleGln 380
 DB 1115 TGCACAAAGGAGCAATGAGACGACTTCTGACAGCTCATGATGATGAGGGCATATCCAG 1174
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
 DB 1175 TATGATATGGCATATGCTGACAAACCTTTCTGCTAAGAAATGAGCTATATGAAGGATTC 1234
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 DB 1235 CATGACAGCTGTGGGAAATCATGCTACTTTCGACGCCACACTTAAGCATTTAAATCC 1294
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440
 DB 1295 ATTGCTCTTGTGCACCCGATTTTCAAGAAGACAAATGAACAGAAATTAACCTTCTGCTC 1354
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460
 DB 1355 AAACAGCACCTCACGATTTGGGACCTGCCATTTACATGTTAGAGAGTGGAGG 1414
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysTrpTrpGluMet 480
 DB 1415 TGGATGCTTTTAAAGGGAAATTTCCCAAGACCAAGTGAATGAAGTGTGTGGAGATG 1474
 QY 481 LysArgGluIleValGlyValGluGluProValProHisAspGluThrTyrCysAspPro 500
 DB 1475 AAGCGAGAGATAGTGGGGGTGGAGACCTGTGCCCATGATGAACATCTGTGACCCC 1534
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheLeuArgTyrThrArgThrLeu 520
 DB 1535 GCATCTCTGTTCCATGTTTGAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1594
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis 540
 DB 1595 TACCAATTCCAGTTTCAAGAACACATTGTTCACAGCAGCTTAATATGAAGGCCCTGTGCAC 1654
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560
 DB 1655 AAATGTGACATCTCAAACTCTACAGAAAGCTGGACAGAAACTGTTCAAATATGCTGAGGCTT 1714
 QY 561 GlyLysSerGluProTrpThrThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetLeu 580
 DB 1715 GGAATAATCAAGACCTTGAGCCCTGAGCATTTGGAATAATGTTGTAGAGCAAAAGCAATGAAT 1774
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600
 DB 1775 GTAAGGCACTGCTCAACTACTTTTGAAGCCCTTTTACTGCTGTGAAGACCAAGCAAG 1834
 QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620
 DB 1835 AATTCCTTTTGTGGATGAGTACCGACTGGAGTCCATATGCAACCAAGCATCAAGT 1894
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluThrPheAsnAspAsnGluMet 640
 DB 1895 AGGATTAAGCCTTAATAATCAGCTTTGAGATGAATGAATGAATGAATGAATGAATGAATGA 1954
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGluThrPheLeuLysValLysAsn 660
 DB 1955 TACCTGTCCGATCATCTGTGTGATATGCTATGAGCAGTACTTTTAAAGTAAATAAT 2014
 QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680

Db 2015 CAGATGATCTTTTGGGGAGGAGATGCGAGTGGCTAATTGAAACCAAGATCTCC 2074
 QY 681 PheAsnPhenValThrAlaProIysAsnValSerAspIleIleProArgThrGluVal 700
 Db 2075 TTTAAATCTTGTTCACACCTGACCTAAAGATGTCTGATATCATCTTCAAGACGTAACTT 2134
 QY 701 GluValAlaIleArgMetSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
 Db 2135 GAAAGGCCATCAGATGTCGCCGAGCCGATCATGATGCTTCCGCTGTAATGACAAAC 2194
 QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
 Db 2195 ACCCTAGAGTTCTCTGGGGATACAGCCAACTGGACCTCTTAACCAAGCCCTGTTTCC 2254
 QY 741 IleTrpLeuIleValPheGlyValIleMetGlyValIleValGlyIleValIleLeu 760
 Db 2255 AATAGCTGATTTGTTTGGAGTTGTGATGGAGTGATAGGTGGGATGTCATCTCTG 2314
 QY 761 IlePheThrGlyIleArgAspArgIysIysAsnIysAlaArgSerGlyGluAsnPro 780
 Db 2315 AACTCAGCTGGATCAGAGATCGGAAGAAATAAAGCAAGAGTGGCAAAATCTT 2374
 QY 781 TyrAlaSerIleAspIleSerIysGlyIysAsnProGlyPheGlnAsnThrAspAsp 800
 Db 2375 TATGCTCCATCGATATTAGCAAGAGAAATAATATCCAGATTTCCAAACACATGATGAT 2434
 QY 801 ValGlnThrSerPhe 805
 Db 2435 GTTCAGACCTCCTTT 2449

RESULT 2

US-10-114-893-85
 : Sequence 85, Application US/10114893
 : Publication No. US20020193567A1
 : GENERAL INFORMATION:
 : APPLICANT: Jacobs, Kenneth
 : APPLICANT: McCoy, John M.
 : APPLICANT: Lavalie, Edward R.
 : APPLICANT: Collins-Racie, Lisa A.
 : APPLICANT: Evans, Cheryl
 : APPLICANT: Meiberg, David
 : APPLICANT: Treacy, Maurice
 : APPLICANT: Bowman, Michael R.
 : APPLICANT: Spaulding, Vikki
 : APPLICANT: Carlin-Duckett, McKeough
 : APPLICANT: Kelleher, Kerry S.
 : APPLICANT: Genetics Institute, Inc.
 : TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES ENCODING THEM
 : FILE REFERENCE: GI 6000-10A
 : CURRENT APPLICATION NUMBER: US/10/114,893
 : EARLIER FILING DATE: 2002-04-02
 : EARLIER APPLICATION NUMBER: 09/413,232
 : NUMBER OF SEQ ID NOS: 321
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 85
 : LENGTH: 3325
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-10-114-893-85

Alignment Scores:
 Pred. NO.: 0 Length: 3325
 Score: 4287.00 Matches: 804
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.88% Mismatches: 0
 Query Match: 99.91% Indels: 0
 Db: 9 Gaps: 0

US-09-978-385-2 (1-805) x US-10-114-893-85 (1-3325)

QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaIleGlnSerThr 20
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 19 ATGTCAAGCTCTTCTTGGCTCTCTTCTCAAGCTTGTCTTAACTGCTCAGTCCACC 78
 QY 21 IleGluGluGlnAlaIysTrpPheLeuAspIysPheAsnIleGluAlaIleAspLeuPhe 40
 Db 79 ATTGAGAACAGGCGCAAGACATTTTGGACAGATTTAACCAAGGCGCAAGACCTGTTTC 138
 QY 41 TyrTrpIleSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluIleValGln 60
 Db 139 TATCAAGGTCACCTTCTTGGAAATTAACACCAATATATCTAGTAGGAATGTCCAA 198
 QY 61 AsnMetAsnAlaGlyAspIysTrpSerAlaPheLeuIysGluGlnSerThrLeuAla 80
 Db 199 AACTGATATATGCTGGGCAAAATGCTGCTTTTAAAGAACAGTCCACACTTGCC 258
 QY 81 GlnMetTrpProLeuGlnGluIleGlnAsnLeuThrValIysLeuGlnLeuGlnAlaLeu 100
 Db 259 CAATGTATCTCAGTACAGAAATTCAGAAATTCACAGTCAAGCTCAGCTGACGCTCTT 318
 QY 101 GlnGlnAsnGlySerSerValIleSerGluAspIysSerIysArgLeuAsnThrIleLeu 120
 Db 319 CAGCAAAATGGGTCTTCTGCTCTCAGAGAACAGCAAGCAAGGTTGAACACAAATTTCA 378
 QY 121 AsnThrMetSerThrIleTyrSerThrGlyIysValCysAsnProAspAsnProGlnGlu 140
 Db 379 AATACATATGAGCACCATCTACATGAGTGAAGAGTTGTATACCAATATATCCACAGAA 438
 QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
 Db 439 TGCTTATTAATCTTGAACCGGTTGAATGAATATGCAACAGTTTATGACTACATATGAG 498
 QY 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyGlnLeuArgProLeuTyr 180
 Db 499 AGGCTCTGGGCTTGGAAACCTGAGATCTGAGAGTGGCAAGCAAGCTGAGGCCATTTAT 558
 QY 181 GluGluIleValValLeuIysAsnGluMetAlaArgAlaAsnIleThrGluAspTyrGly 200
 Db 559 GAAGATATGTGCTTGAAGAAATGAGATGCAAGAGCAAAATCATATATGAGGACTATGG 618
 QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyIysTyrAspTyrArgGly 220
 Db 619 GATTATTGGAGAGAGCATATGAGTAAGTAAGGGGTGATGCTATACATACAGCCCGCCG 678
 QY 221 GlnLeuIleGluAspValGluIleSerThrPheGluGluIleIleAspProLeuTyrGluIleLeu 240
 Db 679 CAGTGTATTACATGATGGAGAACATACCTTTGAAGAGATTAACCATTAATATGACACTCT 738
 QY 241 HisAlaTyrValAlaArgAlaIysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
 Db 739 CATGCTATGTGAGGCAAAATGATGAATGCCTATCTCTATATACATGCAATTTGA 798
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280
 Db 799 TGCTTCCCTGCTCATTTGCTGTGTGATGTGGGTGATGTTGGCAAAATCTGTACTCT 858
 QY 281 LeuThrValProPheGlyGlnIysProAsnIleAspValThrAspAlaMetValAspGln 300
 Db 859 TTGACAGTTCCTTTGGAGCAAGAACCAACCATATGATGATGATGATGATGATGATGATG 918
 QY 301 AlaTrpAspAlaGlnArgIlePheIysGluAlaGluIysPhePheValSerValGlyLeu 320
 Db 919 GCCTGGATGACAGAGAAATATTCAGAGGAGGCCGAGAAAGTTCTTGTATCTTGGTCTT 978
 QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
 Db 979 CTTAATATGCTTAAGGATTTCTGGGAAATTCATGCTTAACGAGACCCAGAAATGTTCAG 1038
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyIysGlyAspPheArgIleLeuMet 360
 Db 1039 AAGGCAAGTGTGACATCCACAGAGTTCGACCTGGGGAAGGAGGAGCATCTGATCTTATG 1098
 QY 361 CysThrIysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380
 Db 1099 TGCACAAAGGTGACATGTAGAGGACTTCTGACAGACTCATGAGATGGGCGCATATCCAG 1158

QY 381 TYRASPMEALATYRALAAGINPROPHLEULEUARGSNGLYALAASNGLUCLYPHE 400
 DB 1159 TATGATATGCAATATCTGACACCTTTTCTGCTAGAGAAAGGACCTATATAGAGATTC 1218
 QY 401 HISGLUALAVALGYLUILEMETSERLEUSERIALAATHRPROLYSHISLEULYSER 420
 DB 1219 CATGAAGCTGTGGGAAATCATGTCTTCTGAGCCACACCTAAGCATTTAAATCC 1278
 QY 421 ILEGLYLEULEUSERPROASPHENGLIUAASPNGLUTHTGLULIASNPHLEULEU 440
 DB 1279 ATGTGCTTCTGTGACCCCATTTTCAAGAACATGAACAAATTAACCTCTGCTCC 1338
 QY 441 LYSGLALALEUTHRIEVALGYTHRLAUPROPHETHRTYRMELEUGLUSTRPARG 460
 DB 1339 AAMCAAGCACTACGATTTGTGGGACTGTGCCATTACTTACATGTAGAGAAAGTGAGG 1398
 QY 461 TRPMETVALPHELYSGLYLUILEPROLYSASPGINTRPMELLYSTRTPGLUMET 480
 DB 1399 TGGATGCTCTTTAAAGGGGAAATCCCAAGACACAGTGGATGAAAAAGTGTGGAGATG 1458
 QY 481 LYSARGGLULIETALGYVALGYLUPROVALPROHISASPGLUHTRYCYASPRO 500
 DB 1459 AAGCGAGAGATAGTGGGTGGTGGAACTGTGCCCATGATGAACATACGTGACCCC 1518
 QY 501 ALASERLEUPHEHISVALSERASNASPTYSERPHELIARGTYTTRHARGTHRLAU 520
 DB 1519 GCATCTCTGTTCATTTCTTAATGATTACTCATTCATTCGATATTACCAAGACCTT 1578
 QY 521 TYRGLNPHENGLNPHENGLUALALEUCYSGLNALAALALYSHISGLYPROLEUHS 540
 DB 1579 TACCAATTCAGATTCAAGAGACATTTGTCAAGCAGCTAAACATGAAGCCCTGCGAC 1638
 QY 541 LYSGLYASPLISERASNSERHGRUALAGLYGLNLYLEULEUPHEASNMETLEUARGLEU 560
 DB 1639 AATGTGACATCTCAACACTTACAGAACTGGACAGAACTGTCAATATGCTGAGGCTT 1698
 QY 561 GLYLYSERGLUPROTTRPHLEUALALEUGLUASNVALVALGYVALYASASNMETASN 580
 DB 1699 GGAATATCAAGAACCTCGAGACCTGAGCATTTGGAATAATGTTGTAGGAGCAAGAACATGAT 1758
 QY 581 VALARGPROLEULANSNTYRPHENGLUPROLEUPHERHTHTRPLEULYSASPLINLYS 600
 DB 1759 GTTAGGCCACTGCTCAACACTTTGAGCCCTTATTACTGCTGTAACCAAGAACGAA 1818
 QY 601 ASNSERPHENGLYTRPSEPTHRASPTRPSERPROTYRALASPGINSERILEULYSVAL 620
 DB 1819 AATTTCTTTTGTGGATGAGTACCGCTGAGTCCATATGCAAGACCAAGCATCAAGTG 1878
 QY 621 ARGLESERLEULYSERIALALEUGLYASPLYSALATYRGLUTRPAASNASPNGLUMET 640
 DB 1879 AGGATTAAGCCTAAATACAGCTTGGAGATTAAGCATATGATGAAGCAATGAATGAA 1938
 QY 641 TYRLEUPHEARGSERSERVALAATYRALAMETARGINTYRPHENULYSVALYSAAN 660
 DB 1939 TACCTGTTCGATCATCTGTGCATATGCTATGAGGCACTACTTTTAAAGTAAAAAT 1998
 QY 661 GLMETILEUPHEGLYGLIUAASPVALARVALAALASNLEULYSPROARGILESER 680
 DB 1999 CAGATGATTTCTTTGGGAGAGAGATGTGCGAGTGGCTAATTTGAACAGAAATCTCC 2058
 QY 681 PHEASNPHEPHEVALTHRALAPOLYSASNVALSERASPILETRPROARGTHRGUVAL 700
 DB 2059 TTTAATTTCTTTGTCACTGACCTAAATGTTCTGATATCATCTTCTGAACTGAAT 2118
 QY 701 GLULYSALALALEARGMETSERARGSERARGILEASNASPALPHENARGLEULASNAPAN 720
 DB 2119 GAAAGGCGCATTCAGAGATGTCGCGAGCCGATCAAGATGCTTCCGTGTGAATGCAAC 2178
 QY 721 SERLEUGLUPHEUGLYLEGINPROTHLEUGLYPROPROKANSGLNPROPROVALSER 740
 DB 2179 AGCCTAGAGATTCTGGGAGATACAGCAACACTTGGAGACCTTCCTAACCAGGCCCTCTTTC 2238

QY 741 ILETRPLEULEVALPHENGLYVALIAMEGLYVALILEVALGYLYLEVALILELEU 760
 DB 2239 ATATGGCGATGTTGTTTGGAGCTGTGATGGAGATGATATGATGTCATTTGCATCCTG 2298
 QY 761 ILEPHERHGRGLIIEARGSPARGLYSLYSASNLYSALARGSERGLYLUASNPRO 780
 DB 2299 ATCTTCACTGGATCGAGATCGAGAGAGAGAAAAATTAACCAAGAACTGGAGAAAACT 2358
 QY 781 TYRALASERILEASPILESERLYSGLYLUASNPNPROGLYPHENGLASNTHRASPNP 800
 DB 2359 TATGCTTCATCATGATTATTAGCAAGAGAAATATATCCAGATTTCAAAACACTGATGAT 2418
 QY 801 VALGINTHRSERPE 805
 DB 2419 GTTCAGACCTCCTTT 2433

RESULT 3
 US-10-028-072-71
 ; Sequence 71, Application US/10028072
 ; Publication No. US20030004311A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Olang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Wacnabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang
 ; TITLE OF INVENTION:
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/028, 072
 ; CURRENT FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/049911
 ; PRIOR FILING DATE: 1997-06-18
 ; PRIOR APPLICATION NUMBER: 60/056974
 ; PRIOR FILING DATE: 1997-08-26
 ; PRIOR APPLICATION NUMBER: 60/059113
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059115
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059117
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059122
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059184
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059263
 ; PRIOR FILING DATE: 1997-09-18
 ; PRIOR APPLICATION NUMBER: 60/059352
 ; PRIOR FILING DATE: 1997-09-19
 ; PRIOR APPLICATION NUMBER: 60/059588
 ; PRIOR FILING DATE: 1997-09-19
 ; PRIOR APPLICATION NUMBER: 60/059836
 ; PRIOR FILING DATE: 1997-09-24
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/062285
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/062287
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/062814
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/062816

1	PRIOR FILING DATE: 1997-10-24
2	PRIOR APPLICATION NUMBER: 60/063045
3	PRIOR FILING DATE: 1997-10-24
4	PRIOR APPLICATION NUMBER: 60/063082
5	PRIOR FILING DATE: 1997-10-31
6	PRIOR APPLICATION NUMBER: 60/063127
7	PRIOR FILING DATE: 1997-10-24
8	PRIOR APPLICATION NUMBER: 60/063327
9	PRIOR FILING DATE: 1997-10-27
10	PRIOR APPLICATION NUMBER: 60/063329
11	PRIOR FILING DATE: 1997-10-27
12	PRIOR APPLICATION NUMBER: 60/063550
13	PRIOR FILING DATE: 1997-10-28
14	PRIOR APPLICATION NUMBER: 60/063561
15	PRIOR FILING DATE: 1997-10-28
16	PRIOR APPLICATION NUMBER: 60/063704
17	PRIOR FILING DATE: 1997-10-29
18	PRIOR APPLICATION NUMBER: 60/063738
19	PRIOR FILING DATE: 1997-10-29
20	PRIOR APPLICATION NUMBER: 60/063755
21	PRIOR FILING DATE: 1997-10-17
22	PRIOR APPLICATION NUMBER: 60/064248
23	PRIOR FILING DATE: 1997-11-03
24	PRIOR APPLICATION NUMBER: 60/064809
25	PRIOR FILING DATE: 1997-11-07
26	PRIOR APPLICATION NUMBER: 60/065186
27	PRIOR FILING DATE: 1997-11-12
28	PRIOR APPLICATION NUMBER: 60/065846
29	PRIOR FILING DATE: 1997-11-17
30	PRIOR APPLICATION NUMBER: 60/066364
31	PRIOR FILING DATE: 1997-11-21
32	PRIOR APPLICATION NUMBER: 60/066453
33	PRIOR FILING DATE: 1997-11-24
34	PRIOR APPLICATION NUMBER: 60/066511
35	PRIOR FILING DATE: 1997-11-24
36	PRIOR APPLICATION NUMBER: 60/066770
37	PRIOR FILING DATE: 1997-11-24
38	PRIOR APPLICATION NUMBER: 60/069212
39	PRIOR FILING DATE: 1997-12-11
40	PRIOR APPLICATION NUMBER: 60/069278
41	PRIOR FILING DATE: 1997-12-11
42	PRIOR APPLICATION NUMBER: 60/069334
43	PRIOR FILING DATE: 1997-12-11
44	PRIOR APPLICATION NUMBER: 60/069694
45	PRIOR FILING DATE: 1997-12-16
46	PRIOR APPLICATION NUMBER: 60/072320
47	PRIOR FILING DATE: 1998-01-23
48	PRIOR APPLICATION NUMBER: 60/073612
49	PRIOR FILING DATE: 1998-02-04
50	PRIOR APPLICATION NUMBER: 60/074086
51	PRIOR FILING DATE: 1998-02-09
52	PRIOR APPLICATION NUMBER: 60/074092
53	PRIOR FILING DATE: 1998-02-09
54	PRIOR APPLICATION NUMBER: 60/077791
55	PRIOR FILING DATE: 1998-03-12
56	PRIOR APPLICATION NUMBER: 60/078910
57	PRIOR FILING DATE: 1998-03-20
58	PRIOR APPLICATION NUMBER: 60/079294
59	PRIOR FILING DATE: 1998-03-25
60	PRIOR APPLICATION NUMBER: 60/079663
61	PRIOR FILING DATE: 1998-02-27
62	PRIOR APPLICATION NUMBER: 60/079728
63	PRIOR FILING DATE: 1998-03-27
64	PRIOR APPLICATION NUMBER: 60/080165
65	PRIOR FILING DATE: 1998-03-31
66	PRIOR APPLICATION NUMBER: 60/081203
67	PRIOR FILING DATE: 1998-04-09
68	PRIOR APPLICATION NUMBER: 60/081229
69	PRIOR FILING DATE: 1998-04-09

1	PRIOR APPLICATION NUMBER: 60/081655
2	PRIOR FILING DATE: 1998-04-14
3	PRIOR APPLICATION NUMBER: 60/081817
4	PRIOR FILING DATE: 1998-04-15
5	PRIOR APPLICATION NUMBER: 60/081818
6	PRIOR FILING DATE: 1998-04-15
7	PRIOR APPLICATION NUMBER: 60/082999
8	PRIOR FILING DATE: 1998-04-24
9	PRIOR APPLICATION NUMBER: 60/083322
10	PRIOR FILING DATE: 1998-04-28
11	PRIOR APPLICATION NUMBER: 60/083545
12	PRIOR FILING DATE: 1998-04-29
13	PRIOR APPLICATION NUMBER: 60/084600
14	PRIOR FILING DATE: 1998-05-07
15	PRIOR APPLICATION NUMBER: 60/084627
16	PRIOR FILING DATE: 1998-05-07
17	PRIOR APPLICATION NUMBER: 60/084637
18	PRIOR FILING DATE: 1998-05-07
19	PRIOR APPLICATION NUMBER: 60/085149
20	PRIOR FILING DATE: 1998-05-12
21	PRIOR APPLICATION NUMBER: 60/085323
22	PRIOR FILING DATE: 1998-05-13
23	PRIOR APPLICATION NUMBER: 60/085338
24	PRIOR FILING DATE: 1998-05-13
25	PRIOR APPLICATION NUMBER: 60/085579
26	PRIOR FILING DATE: 1998-05-15
27	PRIOR APPLICATION NUMBER: 60/085697
28	PRIOR FILING DATE: 1998-05-15
29	PRIOR APPLICATION NUMBER: 60/085704
30	PRIOR FILING DATE: 1998-05-15
31	PRIOR APPLICATION NUMBER: 60/086414
32	PRIOR FILING DATE: 1998-05-22
33	PRIOR APPLICATION NUMBER: 60/086430
34	PRIOR FILING DATE: 1998-05-22
35	PRIOR APPLICATION NUMBER: 60/087106
36	PRIOR FILING DATE: 1998-05-28
37	PRIOR APPLICATION NUMBER: 60/088026
38	PRIOR FILING DATE: 1998-06-04
39	PRIOR APPLICATION NUMBER: 60/088730
40	PRIOR FILING DATE: 1998-06-10
41	PRIOR APPLICATION NUMBER: 60/088741
42	PRIOR FILING DATE: 1998-06-10
43	PRIOR APPLICATION NUMBER: 60/088810
44	PRIOR FILING DATE: 1998-06-10
45	PRIOR APPLICATION NUMBER: 60/088858
46	PRIOR FILING DATE: 19/98-06-11
47	PRIOR APPLICATION NUMBER: 60/089532
48	PRIOR FILING DATE: 1998-06-17
49	PRIOR APPLICATION NUMBER: 60/089599
50	PRIOR FILING DATE: 1998-06-17
51	PRIOR APPLICATION NUMBER: 60/089907
52	PRIOR FILING DATE: 1998-06-18
53	PRIOR APPLICATION NUMBER: 60/089947
54	PRIOR FILING DATE: 1998-06-19
55	PRIOR APPLICATION NUMBER: 60/090349
56	PRIOR FILING DATE: 1998-06-23
57	PRIOR APPLICATION NUMBER: 60/090429
58	PRIOR FILING DATE: 1998-06-24
59	PRIOR APPLICATION NUMBER: 60/090445
60	PRIOR FILING DATE: 1998-06-24
61	PRIOR APPLICATION NUMBER: 60/090538
62	PRIOR FILING DATE: 1998-06-24
63	PRIOR APPLICATION NUMBER: 60/090863
64	PRIOR FILING DATE: 1998-06-26
65	PRIOR APPLICATION NUMBER: 60/091360
66	PRIOR FILING DATE: 1998-07-01
67	PRIOR APPLICATION NUMBER: 60/091519
68	PRIOR FILING DATE: 1998-07-02
69	PRIOR APPLICATION NUMBER: 60/091982
70	PRIOR FILING DATE: 1998-07-07

Alignment Scores:

Pred. No.: 0 Length: 3732
 Score: 4142.00 Matches: 802
 Percent Similarity: 85.33% Conservative: 1
 Best Local Similarity: 85.23% Mismatches: 2
 Query Match: 96.53% Indels: 138
 DB: 9 Gaps: 1

US-09-978-385-2 (1-805) x US-10-028-072-71 (1-3732)

1 MetSerSerSerSerTripleuLeuSerLeuValAlaValAlaThrAlaGlnSerThr 20
 40 ATGTCAAGTCTTCTGCTGCTCTTCTGACCTGTGTGCTGTAACGCTGCTGACGCTGAC 99
 21 IlegluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
 100 ATTGAGGAACAGCCAGCAATTTTGGACAAGTTTAAACCCAGAACCCGAGACCTGTTC 159
 41 TyrGlnSerSerLeuAlaSerThrPasnThrAsnThrAlaThrGluGluAsnValGln 60
 160 TATCAAAAGTCTCTGCTTCTTGGATTATTAACCACTATTTACTGAAAGAGATGTCCAA 219
 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
 220 AACATGAATATATGTGGGACAATGTGCTGCTTTTAAAGAACAGCTCCACACTTGC 279
 81 GlnMetTrpProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuAlaLeu 100
 280 CAAATGTATCCACTACAAAGAAATTCAGATTCACAGTCACGCTTCACTGACGCTCTT 339
 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
 340 CAGCAAAATGGGCTCTTCACTGCTCTCAGAACAGCAAGCAAAAGCTTGAACAACTTCTA 399
 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
 400 AATACATAGACACATCTACAGTACTGTGAAAAGTTTGTAAACCCAGATTAATCCACAA 459
 141 CysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGlu 160
 460 TGGTATTTCTTGAACCAAGCTTGAATGAATGAATGGCAACAGTTTGAACATATGAG 519
 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyr 180
 520 AGGCTCTGGGCTTGGGAAAGCTGGAATCTGAGTCTGGCAACAGCTGAGCCATTATAT 579
 181 GlnGluTrpValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAspTyrGly 200
 580 GAGAGATATGTGTGTTGAAAATGAGATGGCAAGAGCAAAATCATTTATGAGACTATGG 639
 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgly 220
 640 GATTATTTGAGAGAGACATATGAGTAATGGGTAGATGAGTATGACTACAGCGCGGCG 699
 221 GlnLeuIleGluAspValGlnHisThrPheGluGluIleLysProLeuTyrGlnHisLeu 240
 700 CAGTTGATGAGATGTGGAACATCTTGAAGAGATTAACCATTTATATGAACATCTT 759
 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
 760 CAGCTTATGTAGGGCAAGTTGATGATGATCTTCTCTATATGCTCAATTTGGA 819
 261 CysLeuProAlaHisLeuLeuGlnLysPheTrpGlyArgPheTrpThrAsnLeuTyrSer 280
 820 TGGCTCCCTGCTCATTTGCTTGGTGTATGTGGGTAGATTTTGGACAATGTGTACTCT 879
 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
 880 TTGACAGTTCCTCTTGGACAGAAACCAACATAGATGTACTGATGCAATGTGTGACAG 939
 301 AlaTrpAspAlaGlnArgLysPheLysGluAlaGluLysPhePheValSerValGlyLeu 320
 940 GCCTGGATGCAACAGAAATATTCAAGAGCGGAGAAAGTTCTTTGTATCTGTGTCTT 999

321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
 1000 CCTAATATGACTCAAGAGATTTCTGGAAATTTCCATCTTAACGAGACCCAGAGAAATGTCTAG 1059
 341 LysAlaValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
 1060 AAGACATCTGCGCATCCCAACAGCTTGGGAGCTGGGAGAGGGGCACTTCAGATCCCTATG 1119
 361 CysThrLysValThrMetAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380
 1120 TGCACAAAGGTGACAAATGAGACGACTTCTGACAGCTCATCATGATGAGATGGGCGATATCCAG 1179
 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
 1180 TATGATATGGCATATGTGTGCACAACTTTCTCTGTAAGAAATGAGACTTAATGAGACTTC 1239
 401 HisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 1240 CATGAACCTGTGGGAAATCATGTCTCTTTCAGGCCACACCTTAAGCATTTAAATTC 1299
 421 IleglyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440
 1300 ATTGCTTCTGTCCACCGATTTTCAAGAGACAAATGAACAGAAATTAACCTTCTGTCTC 1359
 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460
 1360 AAMCAACACTCAGACATGTGTGGAGCTCTGCCATTTACTTACTGATGTAAGAGATGTGAG 1419
 461 TrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTrpTrpGluMet 480
 1420 TGGATGTCTTTTAAAGGGAAATTTCCCAAGACAGTGGATGTAAGAAATGTGTGGGATG 1479
 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500
 1480 AAGCGAGAGATAGTTGGGGTGGTGGAGCTGTGCCCATATGATTAACATATCTGTGACCC 1539
 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrpThrArgThrLeu 520
 1540 GCATCTCTGTTCCATGTTTCTGATGATTAATCTCATCTGATTAATCAACAAAGACCTTT 1599
 521 TyrGlnPheGlnPheGlnGluAlaLeuGlnGlnAlaAlaLysHisGluGluProLeuHis 540
 1600 TACCAATTCAGATTTCAAGAGCACTTGTCAAGCACATCAACATGAAGGCCCTGTGAC 1659
 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 555
 1660 AATGTGACATCTCAAACTCTACAGAGCTGGACAGAAACTGTT- GTAAGAAATACCTCA 1718
 555 ----- 555
 1719 AAATGTGAACCTCTCTAGTATTCAGTATTAATTCATTTCCATGCTAGGTTTGTATTGG 1778
 555 ----- 555
 1779 ATTTCTTTGTTCTAAGAAATTTTATGAGCTCAAAATGTCTCATTTACAACCAAA 1838
 555 ----- 555
 1839 CATTTAATTTGTGTGACAGAGAACTAGACCATACAAACAAATTGGGTGGCCACCTGT 1898
 555 ----- 555
 1899 TTTCCTCATATCACTACAGCCCTCTTCTGTTGTAATTGGAAGAAAGCGGTTTAA 1958
 555 ----- 555
 1959 GGTGAATATATCTGTTAATATGCAATCTTTCTTATCTGCAGAGCAAAATTTAGCCAA 2018
 555 ----- 555
 2019 GTCAAGAGAGAAACCATATGATCATGATGTAATATATGATGATCTGAGAACCCCTCAA 2078


```

|||||
Db 700 CAGTTGATGAGATGATGACATACCTTGAAGACATTAACCATATATGACATCTT 759
QY 241 HIsAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrTyrIleSerProIleGly 260
Db 760 CAGCTTAATGTGGGCAAGATTGATGATGCTATCTCTCTATATATGATGCAATTTGGA 819
QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheThrAsnLeuTyrSer 280
Db 820 TGGCTCCCTGCTCATTTGGTTGGTATGTGGGAGAGTTTGGCAAAATGTGACTCT 879
QY 281 LeuThrValProPheGlyGlyLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 880 TTGACAGTTCCCTTTGGACAGAAACCAACATAGATGTACTGATGCAATGCTGAGACCAG 939
QY 301 AlaTyrAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320
Db 940 GCGTGGATGACACAGAGAAATTTCAAGAGGCCGAGAAAGTTCTTTGTATCTGTGCTCT 999
QY 321 ProAsnMetThrGlnGlyPheTyrGlyAsnSerMetLeuThrAspProGlyAsnValGln 340
Db 1000 CCTAATATGACTCAAGAGATTCTGGGAAATTTCCATGCTACGAGACCAGAAATGTTGAG 1059
QY 341 LysAlaValCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1060 AAGACGCTGCGCATCCACACCTGGGACCTGGGGAGGCGCACTTCAGATCCTTATG 1119
QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGlnMetGlyHisIleGln 380
Db 1120 TGCACAAAGGTGACAAATGAGACACTTCTGACAGCTCATGATGAGATGGGATATTCAG 1179
QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuAsnArgAsnGlyAlaAsnGlnLysPhe 400
Db 1180 TATGATATGCAATGCTGACACACTTTTCTGTAAGAAATGGAGCTAATGAGAGATTC 1239
QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaIleThrProLysHisLeuLysSer 420
Db 1240 CATGAAAGCTGTGGGAAATCATGTCACCTTCTGACGACACACCTTAACCATTTAAATCC 1299
QY 421 IleGlyLeuLeuSerProAspPheGlnGluAsnGluThrGluIleAsnPheLeuLeu 440
Db 1300 ATTTGCTCTCTGTCACCCGATTTTCAAGAGACAAATGAAACGAATTAATCTCTGCTTC 1359
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTyrArg 460
Db 1360 AAACAAGCACTCAGCATTTGTTGGACTCTGCCATTTACTTACATGTAGAGAGTGGAGG 1419
QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTyrTrpGluMet 480
Db 1420 TGGATGCTCTTTAAAGGGGAAATTTCCAAAGACAGTGGATGAAAAGTGGTGGGAGATG 1479
QY 481 LysATrGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500
Db 1480 AAGCAGAGATAGTTGGGGTGGTGGACCTGTGCCCATGATGAACATACAGTGGACCCC 1539
QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520
Db 1540 GCATCTCTGTTCCATGTTTTCATGATGATTAATCATTCATTCATTAATTAACAAGAGCCCTT 1599
QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaIleLysHisGluGlyProLeuHis 540
Db 1600 TACCAATTCACATTTCAAGAGCACTTGTCAAGCAGCTTAACATGAAAGGCCCTCTGCAC 1659
QY 541 LysCysAspIleSerAsnSerThrGluAlaGlnLysLeuPhe----- 555
Db 1660 AAATGTGACATCTCAAACTTACAGAGCTGACAGAAACTGTT- GTAAGAAATACCTCA 1718
QY 555 ----- 555
Db 1719 AAATGTGAACCTCTCTAGATTACAGATTACTCATTTTCATGCTAGTTTGTATTG 1778
QY 555 ----- 555

|||||
Db 1779 ATTCTTGTCTTAAAGAAATTTTATGGCCTCAAAATGTCTCATTTACAACCAA 1838
QY 555 ----- 555
Db 1839 CATTTAATTTGTGTGACAGACAGAACTAGACCATACAGAAATTTGGTGGCCACTCTT 1898
QY 555 ----- 555
Db 1899 TTTCCTCATATACATACAGACCTCTCTCTCTGTAATTGGAAGAAAGCGGTTTATG 1958
QY 555 ----- 555
Db 1959 GGTGGAATATATCTGTTAATATGCAATCTTTTCTTATCTGCAGAAACAAATTTAGCAA 2018
QY 555 ----- 555
Db 2019 GTCAAGAGAAAGAAACATAGATCATAGATGTAAATATATGTAATCTGGAACCCCTCA 2078
QY 556 ----- 556
Db 2079 AAGCCCTGAAACCCCTTTTGTGTAGCAATATGCTGAGCCTTGGAAATCAGAACCC 2138
QY 566 TrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeu 585
Db 2139 TGGACCTAGCATGGAATAATGTTGTAGAGC- AAGAAATGATGATGAAGCCACTGCTC 2197
QY 586 AsnTyrPheGluProLeuPheThrTyrPheLysAspGlnAsnLysAsnSerPheValGly 605
Db 2198 AACTACTTGTGGCCTTATTTACTGCTGCTAAAGACCAGAAAGAAATCTTTTGGGGA 2257
QY 606 TrpSerThrAspTyrSerProTyrAlaAsp-GlnSerIleLysValArgIleSerLeuLys 625
Db 2258 TGGAGTACCGCATGAGTCCATATGACAGACCAAGATCAAAAGTGAAGATAGCCTTAA 2317
QY 625 SSerAlaLeuGlyAspLysAlaTyrGluTyrPasnAspAsnGluMetTyrLeuPheArgSe 645
Db 2318 ATCAGCTCTTGAGATTAACCATATGATGAATGAAACGACAAATGAATGTACTCTCCGATC 2377
QY 645 rSerValAlaTyrAlaMetArgIleThrPheLeuLysValLysAsnGlnMetIleLeuPhe 665
Db 2378 ATCTGTTGCATATGCTATGAGGAGTACTTTTAAAGTAAATAATCAGATGATCTTTT 2437
QY 665 eGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePhe 685
Db 2438 TGGGAGAGAGATGTGGAGTGGCTAATTTGAAACCAAGATCTCCTTAATTTCTTTGT 2497
QY 685 LThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAlaIleArg 705
Db 2498 CACTGCACCTTAAATAATGTCTGATATCATCTTACAGACTGAGAGTGAAGAGCCATCAG 2557
QY 705 gMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLe 725
Db 2558 GATGTCCTCGGAGCCGTATCATATGATGCTTCCGTGGAAGACAAAGCCTAGATTTCT 2617
QY 725 uGlyIleGlnProThrLeuGlyLysProAsnGlnProProValSerIleGlyLeuIleVal 745
Db 2618 GGGGATACACCAACCACTTGGACCTCTTACACAGCCCCCTGTTCCATATGGCTGATTTGT 2677
QY 745 LpneGlyValAlaMetGlyValIleValValGlyIleValIleLeuIlePheThrGlyI 765
Db 2678 TTTTGGAGTTGTATGAGAGTATGATGTGTTGTCATCTGATCCATCTGATCTCAGTGA 2737
QY 765 eArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnProTyrAlaSerIleAs 785
Db 2738 CAGAGATCGGAGAGAAAGAAATTAAGCAGAGATGAGAAATCTTAAGCCTCATCA 2797
QY 785 pIleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAspValGluThrSerPh 805
Db 2798 TATTAGCAAGAGAAATATATCCAGATTTCCAAAACACTGATGATGTTCAGACTCTT 2857
QY 805 e 805
Db 2858 T 2858

```



```

Db 160 TATCAAGTCTGCTTGGATATTAACCACTATTACTGAAAGACATGTCCAA 219
Qy 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
Db 220 AACATGAATAAATGCTGGGCAAAATGCTGCTTTTAAAGAACAGCTCCACTTCC 279
Qy 81 GlnMetTrpProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuAlaLeu 100
Db 280 CAAATGATCCACTACAGAAATTCAGATTCACAGTCACAGCTTCACCTGAGGCTCTT 339
Qy 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
Db 340 CAGCAAAATGGGCTTTCAGTGTCTCAGAGACAGACAAACGGTTGAAACAATTCCTA 399
Qy 121 AsnThrMetSerThrIleLeuSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
Db 400 AATACAAATGAGCACATCTACAGTACTGAAAGTTTGTAACTCCAGATTAATCCACAAAGAA 459
Qy 141 CysLeuLeuLeuGluProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTrpAsnGlu 160
Db 460 TCTCTTATTAATCTGAACCAAGTTTGAATGAATATGCAACAGTTTATGACTTACCAATGAG 519
Qy 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuLys 180
Db 520 AGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGTGGCAAGCAGCTGAGGCCATTTATAT 579
Qy 181 GluGluTrpValValLeuLysAsnGluMetAlaArgAlaAsnHisTrpGluAspTrpGly 200
Db 580 GAAAGGATATGTGCTTAAAAAATGATGTCACAAAGCAAAATCATTAAGAGACTATGGG 639
Qy 201 AspTrpTrpArgGlyAspTrpGluValAlaAsnGlyValAspGlyTrpAspTrpSerArgGly 220
Db 640 GATTATTTGAGAGAGAGACTATGAATGAATGAGGCTAGTGGCTATGACTTAACCCGCGCC 699
Qy 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTrpGluHisLeu 240
Db 700 CAGTGAATGGAAGATGTGGAACATACCTTTGAAGAGATTAACCATTTATGACAACTCTT 759
Qy 241 HisAlaTrpValArgAlaLysLeuMetAsnAlaTrpProSerTrpLieserProIleGly 260
Db 760 CATGCCCTATGTGAGGGCAAAAGTTGATGATGATGCTTCTCTATATCAGTCCAAATTTGA 819
Qy 261 CysLeuProAlaHisLeuLeuGluLysPheTrpGlyArgPheTrpThrAsnLeuTrpSer 280
Db 820 TGGCTCCCTGCTCATTTCTGTTGATATGTGGGTATATTTTGACAAATCTGTACTCT 879
Qy 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 880 TTGACAGTTCCCTTTGGACAGAAACCAAAACATAGTGTACTGATGCAATGTGTGACACAG 939
Qy 301 AlaTrpAspAlaGlnArgIlePheLysGlnAlaGluLysPhePheValSerValGlyLeu 320
Db 940 GCGCTGGATGACAGACAGATATTTCAAGAGAGCCGAGAAAGTCTTGTATCTGTTGGTCTT 999
Qy 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
Db 1000 CCTATATGACTCAAGAGATCTGGGAAATTCATGCTAAAGCAAGCCAGGAAATGTTCTAG 1059
Qy 341 LysAlaValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1060 AAAGCAGTCTCCATCCACACAGCTTGGAGACCTGGGAAAGGCGCACTCAGATCTTATATG 1119
Qy 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380
Db 1120 TGCACAAAGGTGACATGAGAGACTTCTGTACACTATCTACATGAGATGGGCAATATCCAG 1179
Qy 381 TyrAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
Db 1180 TATGATATGCAATATGCTGACACAACTTTTCTGCTAAGAAATGAGAGCTAATGAAAGATTC 1239
Qy 401 HisGluAlaValAlaGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
|||||
Db 1240 CATACAGCTGTTGGGAAATCATGTCACCTTCTGACCCACACCTTAAGCATTTAAATCC 1299
Qy 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440
Db 1300 ATTGCTCTTGTGTCACCCGATTTTTCAGAGAAAGACATGAACAGAAATAACTTCGTCTC 1359
Qy 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460
Db 1360 AAACAAACACTCACAGATTTGGGACTCTGCTCATTTACTTACATGTTAGGAAGTGGAGG 1419
Qy 461 TrpMetValPheLysGlyGluIleProLysAspGluThrPheLysTrpTrpGluMet 480
Db 1420 TGGATGCTCTTAAAGGGGAAATTTCCAAAGACAGTGGATGAAAGATGGTGGAGATG 1479
Qy 481 LysArgGluIleValAlaGlyValAlaGluProValProHisAspLeuThrTrpCysAspPro 500
Db 1480 AAGCAGAGATATGTTGGGGGTGGTGGAGACCTGTGCCCATGATGAACATCTGTGACCC 1539
Qy 501 AlaSerLeuPheHisValSerAsnAspTrpSerPheIleArgTrpTrpThrArgThrLeu 520
Db 1540 GCATCTCTGTCCATGTTTCTGATGATTAATCTCATTTCTGATATTAACAAAGACCCCTT 1599
Qy 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis 540
Db 1600 TACCAATTCAGATTTCAAGAGACACTTTGTCAAGCAGACTTAACATGAAGGCCCTGTCCAC 1659
Qy 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 555
Db 1660 AAATGTACATCTCAAACTCTACAGAAAGCTGACAGAAACTGTT -GTAAGAAATACCTCA 1718
Qy 555 ----- 555
Db 1719 AAATGTTGAACCTCTCTAGTATTAATCACTATTTCCATGCTAGGTTGATTTGG 1778
Qy 555 ----- 555
Db 1779 ATTTCTTTGTTCTAAAAAGAAATTTTATNGCCTCAAAATGCTCATTTACAAACANA 1838
Qy 555 ----- 555
Db 1839 CATTAAATTTGTGTGACAGACAGAACCTAGACATCAACAATTTGGTGGCCACTCTT 1898
Qy 555 ----- 555
Db 1899 TTCTCCCTATCAATACACAGCCCTCTCTCTGTAATTTGGAAGAAAGCGGTTTATG 1958
Qy 555 ----- 555
Db 1959 GGTGAATATATCTGTTAATATGATCATTTCTTCTATGTCGCAAGCAAAATTTAGCCA 2018
Qy 555 ----- 555
Db 2019 GTCAAAAGAGAAGAACCATAGATCATGATTAATATATGATCAATCTGAAACCCCTCA 2078
Qy 556 ----- 556
Db 2079 AAGGCCCTGACCCCTTTTGTGTGACATATCTGAGCGTTGAAATACAGAACCC 2138
Qy 566 TrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeu 585
Db 2139 TGGACCTTACATGGAATAATGTTGTAGGAGC -AAGAACATGAAATGAGGCACTGCTC 2197
Qy 586 AsnTrpPheGluProLeuPheThrTrpPheLysAspGluAsnLysAsnSerPheValGly 605
Db 2198 AACTACTTGAAGCCCTTATTTACTGCTGGAAGACCAACAAAGAAATCTTTTGGGGA 2257
Qy 606 TrpSerThrAspTrpSerProTrpAlaAsp-GlnSerIleLysValArgIleSerLeuLys 625
Db 2258 TGGAGTACGAGACTGAGAGTCCATATGACAGCCAAACATCAAAAGTGAAGATTAAGCTTAA 2317
Qy 625 sSerAlaLeuGlyAspLysAlaTrpGluTrpAsnAspAsnGluMetLysLeuPheArgSe 645
Db 2318 ATCAGCTCTTGGAGATAAAGCATATGAATGAAGAAAGCAATGAATGATACCTGTCCGATC 2377

```


QY 321 ProasmetInrGInglyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
 Db 1000 CCTAATATAGACTCAAGAGATTCCTGGGAAAATTCATGCTAAACGACGACCAATATGTTCCAG 1059
 QY 341 LysAlaValLysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgLLeuMet 360
 Db 1060 AAGACGCTCTCCATCCACAGCTTGGAGCTGGGAGGAGGCGACTTCAGATCCTTAATG 1119
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGlyMetGlyHisLLeuGln 380
 Db 1120 TGCAACAAGGTCACAAATGAGAGACTTCCTGACAGCTCATCATGAGATGGGCAATATCCAG 1179
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuArgAsnGlyAlaAsnGlyLysPhe 400
 Db 1180 TATGATATAGCATATGCGCAACAACCTTTCTCTAAGAAATGAGCATTAATGAGGATTC 1239
 QY 401 HisGluAlaValGlyGluLLeuMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 Db 1240 CATGAGCTGTGGGAAATCATGTCACTTCTGACGCCACACCTAAGCATTTTAAATTC 1299
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAsnGlnThrGluLLeuAsnPheLeuLeu 440
 Db 1300 ATGGTCTCTGTCACCCGATTTTCAGAAACAAATGAAACAGAAATTAACCTCTCTCTC 1359
 QY 441 LysGlnAlaLeuThrLLeuValGlyThrLeuProPheThrTyrMetLeuGlyLysTrpArg 460
 Db 1360 AAACAGCAGCTCAGCATTTGTTGGGACTCTGCCATTACTTACTTACATGTTAGAGAACTGGAGG 1419
 QY 461 TrpMetValPheLysGlyGluLLeuProLysAspGlnThrPheLysLysTrpTrpGluMet 480
 Db 1420 TGGATGCTCTTAAAGGGAAATTCACAAAGACAGTGAGTGAAGAAAGTGGTGGAGATG 1479
 QY 481 LysArgGluLLeuValGlyValAlaGluProValProHisAspGlnThrTyrCysAspPro 500
 Db 1480 AAGGAGAGATAGTTGGGGTGGTGAACCTGTGCCCATGATGAAACATCTACTGACCC 1539
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheLLeuArgTyrThrArgThrLeu 520
 Db 1540 GCATCTCTGTTCCATGTTCTTGATGATGATTCATTCATTCGATATTCACAAAGGACCTT 1599
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaLysHisGlyGluProLeuHis 540
 Db 1600 TACCAATTCAGATTCAAGAAAGCACTTGTCAACGACCTAAACCTGAAGGCCCTCTGCAC 1659
 QY 541 LysCysAspLLeuSerAsnSerThrGluAlaGlyLysLysLeuPhe 555
 Db 1660 AATGTGACATCTCAACTCTACAGAAAGCTGAGAGAAACTGTGTT-GTAAAGTAACCTCA 1718
 QY 555 555 555
 Db 1719 AATGTGAACCTCTCCTAGATATTCAGATTTACTCATTTCCATGCTAGGTTGTATTTG 1778
 QY 555 555 555
 Db 1779 ATTTCTTTGTTCAAAAAAATTTTATGGCTCAAAATGTCTCATTTTACAACCCAA 1838
 QY 555 555 555
 Db 1839 CATTTAATTTGTGTCAGACAGAGAACTAGACCATACAAATTTGGTGGCCACCTCTT 1898
 QY 555 555 555
 Db 1899 TTCTCCTATCATTAACAGCCCTCTCTCTGTAATTGGAAGAAAGAGGGGTTTAG 1958
 QY 555 555 555
 Db 1959 GGTGGAATATATCTGTTAATATGATCTTTCTTATCTCCAGAAACAAATTTAGCCAA 2018
 QY 555 555 555
 Db 2019 GTCAAAAGAAAGAACCATAGATCATAGTAAATATATGTACATCTGGAACCCCTCAA 2078
 QY 556 -----AsnMetLeuArgLeuGlyLysSerGluPro 565

Db 2079 AAGGCCCTGACCCCTTTTGTGTAGCAATATGCTGAGGCTTGGAATTCAGAACCC 2138
 QY 566 TrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeu 585
 Db 2139 TGGACCTATACATTTGAAAATGTTGTAGAGGC-AAAGACATATAAATGAGGCACTGCTC 2197
 QY 586 AsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAspLysAsnSerPheValGly 605
 Db 2198 AACTACTTTGAGCCCTTATTTACTGCTGCTGMAAGCCACAAACAGATTTCTTTGTGGCA 2257
 QY 606 TrpSerThrAspTrpSerProTyrAlaAsp-GlnSerLLeuValArgLLeuSerLeuLys 625
 Db 2258 TGGAGTACGAGCTGAGTGCATATGACAGACCCAAAGCATCAAAAGTGAAGATTAAGCTTAA 2317
 QY 625 sSerAlaLeuGlyAspLysAlaTyrLLeuTrpAsnAspAsnGluMetTyrLeuPheArgSe 645
 Db 2318 ATCAGCTCTGGAGATTAAGCATATGAAATGAAACGACAAATGAAATGTACTCTTCGATC 2377
 QY 645 rSerValAlaTyrAlaMetArgGlnThrPheLeuLysValLysAsnGluMetLLeuPhe 665
 Db 2378 ATCTGTTGCATATGCTATGAGCAGTACTTTTAAAGTAAAAATCAGATATTTCTTTT 2437
 QY 665 eGlyGluGluAspValArgValAlaAsnLeuLysProArgLLeuSerPheAsnPhePheVal 685
 Db 2438 TGGGAGAGAGATGTGCGAGTGCCTAATTTGAACCAAGATCTCCTTAATTTCTTGT 2497
 QY 685 LThrAlaProLysAsnValSerAspLLeuLeuProArgThrGluValGlyLysAlaLLeuArg 705
 Db 2498 CACTGCACCTTAAATATGTCTGATATCATTCCTAGAACGTGAAGTGAAGGACCATCAG 2557
 QY 705 gMetSerArgSerArgLLeuAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLe 725
 Db 2558 GATGTCCCGAGCGCATATGATGATCTTCCGCTCGAATGACAAACAGCTGAGATTTCT 2617
 QY 725 uGlyLLeuProThrLeuGlyProProAsnGlnProProValSerLLeuTrpLeuLLeuVal 745
 Db 2618 GGGGATACAGCCAACTGTGACCTCTCAACAGCCCTGTTCCATATGCTGATGTTGT 2677
 QY 745 lPheGlyValValMetGlyValLLeuValGlyLLeuLLeuLLeuLLeuPheThrGlyL 765
 Db 2678 TTTTGGATGTGATGGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2737
 QY 765 eArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnProTyrAlaSerLLeuAs 785
 Db 2738 CAGAGTTCGAGACAGAAATTAAGCAAGAGTGGAGAAAAATCTTATGCTCCATCGA 2797
 QY 785 pLLeuSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPhe 805
 Db 2798 TATTTACCAAGAGGAAATTAATCCAGGATTTCCAAAACACTGATGATGTTCAACACTCCTT 2857
 QY 805 e 805
 Db 2858 T 2858
 RESULT 8
 US-10-176-918-71
 ; Sequence 71, Application US/10176918
 ; Publication No. US20030027275A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel
 APPLICANT: Matanabe, Colin K
 APPLICANT: Wood, William
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P333081C382
 CURRENT APPLICATION NUMBER: US/10/176,918
 PRIORITY FILING DATE: 2002-06-20
 Prior Application removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 71
 LENGTH: 3732
 TYPE: DNA
 ORGANISM: Homo Sapien
 US-10-176-918-71

Alignment Scores:

Pred. No.: 0 Length: 3732
 Score: 4142.00 Matches: 802
 Percent Similarity: 85.338 Conservative: 1
 Best Local Similarity: 85.238 Mismatches: 2
 Query Match: 96.538 Indels: 138
 Gaps: 1

us-09-978-385-2 (1-805) x us-10-176-918-71 (1-3732)

QY 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
 DB 40 ATGTCAAGCTCTTCCGCTCTCTCCACCTGTGTGTAAGCTGCTCAGTCCACC 99
 QY 21 ILeGluGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
 DB 100 ATTAGAGAACAGGCCAAGACATTTTGGACAAGTTTACCACAGAACCCAAACCTGTTC 159
 QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTrpAsnIleThrGluGluAsnValGln 60
 DB 160 TATCAAAAGTTCACCTCTCTTGGAAATTTAAACACCAATATTACGAAGAAATGTCCA 219
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAla 80
 DB 220 AACATAAATATGCTGGGACAATGTGTGCTTTTAAAGAACAGTCCACACTGTGC 279
 QY 81 GlnMetTrpProLeuGlnIleGlnAsnLeuThrValLysLeuGlnIleAlaLeu 100
 DB 280 CAAATGTATCCACTACAAATAATTCGATCTCACAGTCAAGCTTCAGCTGAGGCTCT 339
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
 DB 340 CACCAAAATGGCTCTCACTGCTCTCAGAACAGAACGAAACGGTTGAACACAATTTCTA 399
 QY 121 AsnTrpMetSerThrIleLeuTrpSerThrLysValCysAsnProAspAsnProGlnGlu 140
 DB 400 AATACATATGACCACTACACTACTGAGAAAGTTGTAAACCCAGATATATCCACAAGAA 459
 QY 141 CysIleuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
 DB 460 TCGTATTACTGTAACCAAGTTTGAATGAATAATGCAACAGTTTAACTATACATGAG 519
 QY 161 ArgLeuTrpAlaTrpLeuSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTrp 180
 DB 520 AGCCTCTGGGCTTGGAAAGCTGGAGATCTGAGTGGCAACAGCTGAGGCATTAATAT 579
 QY 181 GlnGluTrpValValLeuLysAsnGlnMetAlaArgAlaAsnHisTrpGluAspTyrGly 200
 DB 580 GAAAGAGATGTGCTTGAATAATGATGAGCAAGCAAAATCATTTATGAGAGATAGGG 639
 QY 201 AspTyrTrpArgLysAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgLys 220
 DB 640 GATTATTGGAGAGACATATAGTAATGAGGTGATGAGTATGATACAGCGCGGCG 699
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTrpGluHisLeu 240

DB 700 CAGTTGATGGAAGATGTGAACATACCTTTGAAGAGATTAACCATATTATGAACATCTT 759
 QY 241 HisAlaTyrValaArgAlaLysLeuMetAsnAlaTyrProSerLysIleSerProIleGly 260
 DB 760 CATCCCTATGAGAGGCAAGTGTATGAATGCCATCTCTCCATATACATCACTCAATTTGA 819
 QY 261 CysLeuProAlaHisLeuLeuGluLysPheMetTrpGlyLysTrpPheTrpPheAsnLeuTrpSer 280
 DB 820 TGCCTCCCTGCTCATTTGCTTGGTATGATGTGGGGATGATTTTGGACAAATCTGTACTCT 879
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
 DB 880 TTGACAGTCCCTTGGACAGAAACCAACATAGATGTACTGATGCAATGAGTGAGACAG 939
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320
 DB 940 GCCGTGGATGCAAGAAATATTCAAAGAGGCCAGAAAGTTCTTGTATCTGTTGGCTCT 999
 QY 321 ProAsnMetTrpGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
 DB 1000 CCTATATGACTCAAGGATTTCTGGGAAATTTCCATGCTAACGAGCCAGAAATGTTTCAG 1059
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
 DB 1060 AAACGACGTCCCATCCACAGCTTGGGACCTGGGGAAGGCGACATTCAGGATCTTATG 1119
 QY 361 CysThrLysValIleThrMetAspPheLeuThrAlaHisHisGlnMetGlyHisIleGln 380
 DB 1120 TGCACAAAGGTGACAAATGAGACGACTCTTGACACTATATGATGAGGGGCAATTCAG 1179
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
 DB 1180 TATGATATGCGATATGCTGCGCAACCTTTCTGCTAGAAATGAGAGTATGAAGATTC 1239
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 DB 1240 CATGAAGCTGTGGGGAATCATCTCACTTCTGACCCACACCTTGAAGCATTTAAATTC 1299
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440
 DB 1300 ATTGCTCTCTGTCACCCGATTTTCAGAGACATATGAACAGAAATAACTCTGCTC 1359
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTrpMetLeuGluLysTrpArg 460
 DB 1360 AAACAGACACTCAGGATTTGGGAGCTGCGACTGCACTTACTTACATGTTAGAGAGTGAGG 1419
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGluTrpMetLysLysTrpTrpGluMet 480
 DB 1420 TGGATGCTCTTAAAGGGGAATTCACCAAGACAGTGGATGAATAAGTGTGGAGATG 1479
 QY 481 LysArgGluIleValGlyValGluIleProValProHisAspGluThrTrpCysAspPro 500
 DB 1480 AAGCAGAGATATGTTGGGCTGTGGAACTGTGCCCATGATGAAACATATCTGACCCC 1539
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520
 DB 1540 GCATCTGCTGTCATGTTCTGATGATATACCATTTCTGATATATACACAGAGACCTT 1599
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnIleAlaLysHisGluGluProLeuHis 540
 DB 1600 TACCAATTCAGTTTCAAGAGCACTTGTTCACACACCTAAACATGAAAGGCCCTGTGAC 1659
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyLysLysLeuPhe 555
 DB 1660 AATGTGACATCTCAAACTCTACAGAAAGCTGACAGAAACTGTT -GTAAACAAATACCTCA 1718
 QY 555 555
 DB 1719 AATGTGAACTCTCTAGTATTCAGTATTAATTAATTCATTCAGGCTAGGTTGATTTG 1778
 QY 555 555
 DB 1779 ATTCTCTTTGTTCTAAAGAAATTTATGGCTCAAAATGTCTCTCATTTTACAACCAAA 1838

	RESULT 9	
US-10-176-921-71	/ Sequence 71, Application US/10176921	
	/ Publication No. US20030027276A1	
	/ GENERAL INFORMATION:	
/ APPLICANT:	Baker, Kevin P.	
/ APPLICANT:	Beresini, Maureen	
/ APPLICANT:	DeForge, Laura	
/ APPLICANT:	Desnoyers, Luc	
/ APPLICANT:	Filvaroff, Ellen	
/ APPLICANT:	Gao, Wei-Qiang	
/ APPLICANT:	Gerritsen, Mary E.	
/ APPLICANT:	Goddard, Audrey	
/ APPLICANT:	Godowski, Paul J.	
/ APPLICANT:	Gurney, Austin L.	
/ APPLICANT:	Sherwood, Steven	
/ APPLICANT:	Smith, Victoria	
/ APPLICANT:	Stewart, Timothy A.	
/ APPLICANT:	Tumas, Daniel	
/ APPLICANT:	Watanabe, Colin K	
/ APPLICANT:	Wood, William	
/ APPLICANT:	Zhang, Zemin	
/ TITLE OF INVENTION:	ACIDS ENCODING THE SAME	
/ FILE REFERENCE:	P330R1C28	
/ CURRENT APPLICATION NUMBER:	US/10/176,921	
/ PRIOR FILING DATE:	2002-06-20	
/ PRIOR Application removed - See File Wrapper or Palm		
/ NUMBER OF SEQ ID NOS: 550		
/ SEQ ID NO 71		
/ LENGTH: 3732		
/ TYPE: DNA		
/ ORGANISM: Homo Sapien		
US-10-176-921-71		
Alignment Scores:		
Pred. No.: 0	Length: 3732	
Score: 4142.00	Matches: 802	
Percent Similarity: 85.33%	Conservative: 1	
Best local Similarity: 85.23%	Mismatches: 2	
Query Match: 96.53%	Indels: 138	
DB: 9	Gaps: 1	
US-09-978-385-2 (1-805) x US-10-176-921-71 (1-3732)		
OY 1 MetSerSerSerSerTrpLeuLeuLeuSerLeuValAlaValThrAlaIagInSerThr 20		
Db 40 ATGTCAACTCTTCTCGGCTCTCTCTCAAGCTTTTGCTGAAGCTGCTCACTCCACC 99		
OY 21 IlegIugIuAlaLalVstFthPhelenuAsplVsPheasnHtsgIuaIagiAsPluePhe 40		
Db 100 ATTGGAGCACAGGCCAGACATTTTTGGACAAGTTTAACCCAGCAAGCCGAAGACCTGTTC 159		
OY 41 TyrGInserSerLeuAlaserTrpAsnTyrsophrnsInlempIrngIuAsnVaGlIn 60		
Db 160 TATCAAAGTACTCTTCTCTTCTGGAAATTATACACCAATATATACTGAAGAAGAAATGTCOA 219		
OY 61 AsmetAsnaSnaIaglAsPlVsTrpSerAlaPheLeuysGIugInSerThrLeuAla 80		
Db 220 AACCATGAATATATGCTGGGGGACAAATATGGCTGCTTTTAAAGAGACAGTCCACACTTCCC 279		
OY 81 GlnMetYrrProLeuGIngluIllegInasneuthrVallylsueGInleuGInAlaleu 1000		
Db 280 CAAATGATCCATCAACAAGAAATTCAGAAATCTCAACTCAAGCTTCAGCTGCAGGCTCTT 339		
OY 101 GlnGInsnngIySerSerValLeuSerGIAsPlVsSerLySrIguLeuAsnThrIleLeu 120		
Db 340 CAGCAAAATGGSGTTCAGTCTCTCTCAAGAACAGACAGCAAGCGTTGAACACAAATTCTA 399		
OY 121 AsnThrMetSerThrIleyrSerThrgILyIsValCyAsnBroAspAsnProGInglu 140		
Db 400 AAATCAATATGACACCAATCTACAGTACGTGAAGATTTGTAAACCAGATATATCCAGAA 459		

QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluMetAlaAsnSerLeuAspTyrAsnGlu 160
 |||||
 Db 460 TGCCTATTATCTGAACCGAGTTGATGAATAATGCGAACAGTTTGACATACATAG 519
 QY 161 ArgLeuTrpAlaTrpGlySerTrpArgSerGluValGlyIysGlnLeuArgProLeuTyr 180
 |||||
 Db 520 AGGCTCTGGGCTGGGAAAGCTGGAGATCTGAGTCGGCAAGCAGCTAGGCGCATATAT 579
 QY 181 GluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGlyAspTyrGly 200
 |||||
 Db 580 GAAGGATAGTGTCTTGAATAATGAGATGCGCAAGCAAAATCATTAATGAGAGCTATGG 639
 QY 201 AspTyrTrpArgGlyAspTyrGlyValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
 |||||
 Db 640 GATTATTGAGAGAGACATATGAGTAATGGGTAGATGGCTATGACTACAGCCGGCG 699
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240
 |||||
 Db 700 CAGTTGATTGAGAGATCTGGACATACCTTTGAAGAGATTAACCATTAATGACATCTT 759
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
 |||||
 Db 760 CATGCTATGTAGAGGCAAAAGTTGATGAATGCTTCTCTATATCAGTCCAAATGTGA 819
 QY 261 CysLeuProAlaHisLeuLeuGlnLysPheMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280
 |||||
 Db 820 TGCCTCCCTGCTCATTTGCTTGGTATATGTGGGTAGATTTTGGCAAAATGTACTCT 879
 QY 281 LeuThrValArgPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
 |||||
 Db 880 TTGACAGTTCCTTTGAGACAGAAACCAACATGATGTACTGATGCAATGGTGGACAG 939
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320
 |||||
 Db 940 GCGTGGATGACACAGAAATATTCAGAGAGCGGAGAGATCTTTGTATCTTTGGCTT 999
 QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
 |||||
 Db 1000 CCTAATATGATGCTCAAGAGATTTGGGAAATTCATGCTTAAGCGACCCAGAAATGTG 1059
 QY 341 LysAlaValAlaCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
 |||||
 Db 1060 AAAGCAGTGTGCTCCACAGCTGGGACCTGGGGAAAGGGCGACTCCGATCTCTTAG 1119
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisIleGluMetLysIleGln 380
 |||||
 Db 1120 TGCACAAAGGTGACAAATGAGCACTTCTGACAGCTCATCATGAGATGGGCGATATCC 1179
 QY 381 TyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
 |||||
 Db 1180 TATGATATGAGCATATGCTGCACAACTTTTCTGCTAAGAAATGAGACTAATGAGATTC 1239
 QY 401 HisGluAlaValGlyIleLeuMetSerLeuSerAlaIleThrProLysHisLeuLysSer 420
 |||||
 Db 1240 CATGAAGCTGTGGGAAATCATGTCACCTTTGACAGCACACCTAAGCATTTAAATCC 1299
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440
 |||||
 Db 1300 ATTGGTCTTCTGTCACCCGATTTTCAAGAGACAAATGAAACGAATTAATCTTCGCTC 1359
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460
 |||||
 Db 1360 AAACAAGCACTCAGATGTGTGGACTCTGCCATTACTTACTATGTTAAGAGAGTGGAG 1419
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrpGluMet 480
 |||||
 Db 1420 TGGATGCTTTTAAAGGGGAAATTTCCAAAGACAGTGTATGATAAAAGGTGGGAGAG 1479
 QY 481 LysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCysAspPro 500
 |||||
 Db 1480 AACCGAGAGATATGTGGGGGTGGGAACCTGTGCCCATGATGATAAACAATACGTGTGAC 1539
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520

|||||
 Db 1540 GCATCTGTTTCATGTTTCTGATGATTAATCTTCAATTCATGATATTAACAAAGACCTT 1599
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis 540
 |||||
 Db 1600 TACCAATTCAGATTTCAGAAAGCACTTTGTCAAGCAGCTTAACATGAAAGCCCTGTGAC 1659
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 555
 |||||
 Db 1660 AAATGTACATCTCAAACTCTACAGAAAGCTGACAGAAAGTGT-GTAAAGAAATACCTCA 1718
 QY 555 ----- 555
 Db 1719 AAATGTGAACCTCTCTAGTATTAAGTATTAATCAATTCATCCAGCTAGGTTTGTATTG 1778
 QY 555 ----- 555
 Db 1779 ATTCTTTGTTCTAAGAAATTTTATGGCTCAAAATGTCTCTATTACAAACCAAA 1838
 QY 555 ----- 555
 Db 1839 CATTTAATTTGTGTGTCAGACAGAACCTAGACCATACAACAATTTGGGTGGCCACCTTT 1898
 QY 555 ----- 555
 Db 1899 TTCTCCTATCATTAACACAGCCCTCTCTCTGTGTAATTGGAAGAAAGAGCGGTTAG 1958
 QY 555 ----- 555
 Db 1959 GGTGAATATATCTGTTAATATGCAATCTTTCTTATCTCCAGAAACCAATTTAGCCAA 2018
 QY 555 ----- 555
 Db 2019 GTCAAAAGAGAAACCATAGATCATATATATATGATCATCTGAGAACCCCTCAA 2078
 QY 556 ----- 556
 Db 2079 AAGGCCCTGAACCCCTTTTGTGTAGCAATATGTGAGGCTGGAAATACAGAACCC 2138
 QY 556 ----- 556
 |||||
 Db 566 TrpThrLeuAlaLeuGluAsnValValGlyValAlaLysAsnMetAsnValArgProLeuLeu 585
 |||||
 Db 2139 TGGACCTCAGATTTGGAAATATGTTGTAGAGAC-AAAGAACATGATATAGGCCCATCTCTC 2197
 QY 586 AsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGly 605
 |||||
 Db 2198 AACTACTTTGAGCCCTTATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2257
 QY 606 TrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSerLeuLys 625
 |||||
 Db 2258 TGGAGTACCGAGCTGAGTCCATATGACAGACCCCAAGCATCAAAAGTGAAGATTAAGCTTAA 2317
 QY 625 SSerAlaLeuGlnLysAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPheArgSe 645
 |||||
 Db 2318 ATCAGCTTTTGGAGATTAAGCATATGATGAGACAGACATGAATATGACTGTTCCGATC 2377
 QY 645 rSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetIleLeuPhe 665
 |||||
 Db 2378 ATCTGTTCATATGCTATGAGGAGTACTTTTAAAGTAAATAATCAGATGATCTTTT 2437
 QY 665 eGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePheVal 685
 |||||
 Db 2438 TGGGAGGAGAGATGTGCGAGTGGCTAAATTTGAAACCAAGATTCCTTAAATTTCTTTGT 2497
 QY 685 lThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAlaIleArg 705
 |||||
 Db 2498 CACTGCACCTAAATATGCTGTGATATCTCTAAGACGAACTGAATGAAGGCGCATCAG 2557
 QY 705 GMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLe 725
 |||||
 Db 2558 GATGTCCCGGAGCCCTATCAATGATGATGCTTCCGTGTGATGACAAACGCTAGAGTTTCT 2617
 QY 725 uGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerIleTrpLeuIleVal 745
 |||||

Db 2618 GGGGATACAGCAACACTTGACCTCTTAACAGCCCGCTTTCATATGCTGATGT 2677
 QY 745 IPhEGYVALIValMetGlyValIleValValGlyIleValIleLeuIlePheThrGlyI 765
 Db 2678 TTTTGGAGTGTGATGGAGTGTAGTGTGATGCTTTCATCCGATCTTCACGGGAT 2737
 QY 765 eATGASPARGLYSLYSAsnLYSAlaArgSerGlyLYSAsnProTYRAlSerIleAs 785
 Db 2738 CAGAGATCCGAGACAAATAATTAACCAAGAGAGTGAAGAAATCCTTATGCTCCATGCA 2797
 QY 785 pIleSerLYSGLYSAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPh 805
 Db 2798 TATTAACAAGAGAAATAATATCCAGATTCACAAACACTGATGATGCTCAGACCTCTT 2857
 QY 805 e 805
 Db 2858 T 2858

RESULT 10

US-10-137-865-71
 ; Sequence 71, Application US/10137865
 ; Publication No. US20030032155A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerlitsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P330R1C154
 ; CURRENT APPLICATION NUMBER: US/10/137,865
 ; PRIORITY FILING DATE: 2002-05-03
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 71
 ; LENGTH: 3732
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-137-865-71

Alignment Scores:

Pred. No.: 0 Length: 3732
 Score: 4142.00 Matches: 802
 Percent Similarity: 85.33% Conservative: 1
 Best Local Similarity: 85.23% Mismatches: 2
 Query Match: 96.53% Indels: 138
 DB: 9 Gaps: 1

US-09-978-385-2 (1-805) x US-10-137-865-71 (1-3732)

QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
 Db 40 ATGTCAAGCTCTTCTGCTCTCTCTCAAGCTTGTCTCTTAAGCTGCTCACTCCACC 99
 QY 21 ILeGluGluGlnAlaValThrPheLeuAspLYSAsnHisGluAlaGlnAspLeuPhe 40
 Db 100 ATTAGAGCAAGCGCAAGACATTTTGGACAGATTAAACCAAGCGCAAGACCTGTTTC 159
 QY 41 TyGlnSerSerLeuAlaSerTrpAsnTrpAsnThrAsnIleThrGluGluAsnValGln 60

Db 160 TATCAAAATGCTTCCTTCTTGGATTAATTAACCAATATTAATCAAGAGATGTCGAA 219
 QY 61 AsnMetAsnAsnAlaGlyAspLYSTripSerAlaPheLeuLYSGluGlnSerThrLeuAla 80
 Db 220 AACATGAATTAATGCTGGGGACAAATAGTCTCTTTTAAAGAACAGTCCACACTTGCC 279
 QY 81 GlnMetLYRProLeuGlnGluIleGlnAsnLeuThrValLYSLeuGlnLeuGlnAlaLeu 100
 Db 280 CAAATGATTCACATCAAGAAATTCAGATCTCAGACGACAGCTTCAGCTGACGCTCTT 339
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLYSTripSerLYSArgLeuAsnThrIleLeu 120
 Db 340 CAGCAAAATGGGCTTTCAGTCTCTCTCAGAAAGACAAAGCAAGCGTTGAACAAATCTA 399
 QY 121 AsnThrMetSerThrIleTYRSerThrGlyLYSValLYSAsnProAspAsnProGlnGlu 140
 Db 400 AATACAAATGACACCATCTACAGTACGTGAAGATTTGTAAACCAAGATATCCCAAGAA 459
 QY 141 CysLeuLeuLeuGluProGlyLYSAsnGluIleMetAlaAsnSerLeuAspTYRAsnGlu 160
 Db 460 TGCTTATTACTTGAACCAAGCTTTGATGAATTAATGGCAAAACAGTTTAAGACTACATGAG 519
 QY 161 ArgLeuTrpPalaTrpGluSerTrpArgSerGluValGlyLYSGlnLeuArgProLeuTYR 180
 Db 520 AGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGCTGGCAAGCTGAGGCTGATATAT 579
 QY 181 GluGluIuTYRValValLeuLYSAsnGluMetAlaArgAlaAsnHisTYRGLuAspTYRGLY 200
 Db 580 GAAGAGTATGTGCTCTTAATAAATGATGCGACAGCAAAATCATTTATGAGGACTATGGG 639
 QY 201 AspTYRTrpArgGlyAspTYRGLYValAsnGlyValAspGlyTYRAspTYRSerArgGly 220
 Db 640 GATTAATGAGAGAGGAGACTATGAATTAATGGGAGATGCTATGACTACAGCGCGCG 699
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLYSProLeuTYRGLuHisLeu 240
 Db 700 CAGTTGATGAAGATGTGAACATCTCTTGAAGATTAACCATTAATTAAGACATCTT 759
 QY 241 HisAlaIuTYRValArgAlaLYSLeuMetAsnAlaTYRProSerTYRILESerProIleGly 260
 Db 760 CATGCCATATGTAGAGGCAAGTGTGAATGCTCTCTCTATATCAGTCCATATGCA 819
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTYRSer 280
 Db 820 TCCGACCCCGCTCAATTTGCTGTGTATGTGGGGTAGATTGGAACAAATCTGACTCT 879
 QY 281 LeuThrValProPheGlyGlnLYSProAsnIleAspValThrAspAlaMetValAspGln 300
 Db 880 TTGACAGTTCCTCTTGGACAGAAACCAACATGATGTACTGATGACATGATGTGACCG 939
 QY 301 AlaTrpAspAlaGlnArgIlePheLYSGluAlaGluLYSPhePheValSerValGlyLeu 320
 Db 940 GCTTGAGATGCACAGAGATATTTCAAGAGGCGCAAGATTCTTGTATCTGTGGCTCTT 999
 QY 321 ProAsnMetThrGlnGlyPheTrpLeuAsnSerMetLeuThrAspProGlyAsnValGln 340
 Db 1000 CCTAATATGACCAAGAGATTCGCGAAATTCATCAGCTCAACGACCGACGAAATGTTAG 1059
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLYSAspPheArgIleLeuMet 360
 Db 1060 AAAAGCATGTGCTATCCACACACTTGAGCTGGGAGGAGGCGCATTTAGATCTCTTATG 1119
 QY 361 CysThrLYSValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380
 Db 1120 TGCACAAAGTGTACATGAGACATCTTCGACAGCTCATCATGATGAGATGGGATATCCAG 1179
 QY 381 TYRAspMetAlaTYRAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
 Db 1180 TATGATATGCAATATGCTGACACACCTTTTCTGCTAAGAAAGAGCAATTAAGAGATTTC 1239
 QY 401 HisGluAlaValGlyIuIleMetSerLeuSerAlaAlaThrProLYSHisLeuLYSser 420
 Db 1240 CATGAAGCTGTGGGAAATCATGTCATCTTCTGACGACCACTTAAGCATTTAAATCC 1299

Score: 4142.00 Matches: 802
 Percent Similarity: 85.33% Conservative: 1
 Best Local Similarity: 85.23% Mismatches: 2
 Query Match: 96.53% Indels: 138
 DB: 9 Gaps: 1

US-09-978-385-2 (1-805) x US-10-140-474-71 (1-3732)

QY 1 MetSerSerSerSerTrpLeuLeuLeuSerLeuValAlaValThrAlaIleGlnSerThr 20
 DB 40 ATGCAAGCTCTCTCCCTGGCTCTCTGAGCTTGTCTGCTTAACCTGCTGCTGCTCACC 99
 QY 21 IleGlnIleGlnAlaIleValThrPheLeuAspLysPheAsnIleGlnIleGlnAspLeuPhe 40
 DB 100 ATTGAGAACAGGCCAAGACATTTTGGACAGTTTACCACGACGCGAAGACCTGCTTC 159
 QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGlnIleGlnValGln 60
 DB 160 TATCAAGTTCTACTGCTTGTGGAAATTAACACCAATATTACTGAAAGAAATGTCCAA 219
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuIleGlnIleGlnSerThrLeuAla 80
 DB 220 AACATGATATATGCTGGGCAAAATGCTGCTTTTAAAGAACGTCACACTTGGC 279
 QY 81 GlnMetTyrProLeuGlnIleGlnAsnLeuThrValIleGlnIleGlnIleGlnAlaLeu 100
 DB 280 CAATGTATCCACTACAGAAATTCAGAAATCTCAGCTCAGCTGCGACGCTCTT 339
 QY 101 GlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThrIleLeu 120
 DB 340 CAGCAAAATGGCTCTCAGCTCTCTCAGAAAGACAGCAAGCGTGAACACAAATCTTA 399
 QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValIleValIleValIleValIleValIleVal 140
 DB 400 AATCAATGAGCACCATCTCAGTACTGGAAGAAATTGTAAACCCAGTAATTCACAGAA 459
 QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGln 160
 DB 460 TGCATTACTTGACACAGGTTTGAATGAATATATGCAACAGCTTAAAGCTACAAATGAG 519
 QY 161 ArgLeuTrpAlaTrpGlnSerTrpArgSerGlnValGlyLysGlnLeuArgProLeuTyr 180
 DB 520 AGGCTCTGGGCTGGGAAGCTGGAGATCTGAGGCTCGGCAAGCAGCTGAGGCATTTAT 579
 QY 181 GlnIleIleValIleValIleValAsnGlnMetAlaArgAlaAsnIleIleIleIleIleIleIle 200
 DB 580 GAAAGATATGTTGCTTGAAGAAATGAGATGCAAGACCAATTAATTAAGAGCTATGGG 639
 QY 201 AspTyrTrpArgGlyAspTyrGlnValAlaAsnGlyValAlaAspLysTyrAspTyrSerArgGly 220
 DB 640 GATATATTGAGAGAGAGACTATGAAGTAATGGGGTATGCTATGCTATGCTACAGCGCGGC 639
 QY 221 GlnLeuIleGlnAspValGlnHisThrPheGlnIleIleLysProLeuTyrGlnHisLeu 240
 DB 700 CAGTGTGATTGAAGATGGGAACATACCTTGAAGAGATTAACCATTAATTAAGAACTTT 759
 QY 241 HisAlaTyrValAlaArgAlaIleValIleValMetAsnAlaTyrProSerTyrIleSerProIleGly 260
 DB 760 CATCCCTATATGAGGAGCAAAATGATGAATGCCATCTCTATATCAATCAATGGA 819
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280
 DB 820 TGCCTCCCTGCTCATTTGCTTGTGATATATGGGGTATGATTTTGGACAAATCTGTTCTT 879
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
 DB 880 TTGACAGTCTCCCTTGGAGCAAGAAACCAACATAGATCTACTGATGCAATGCGGACAG 939
 QY 301 AlaTrpAspAlaGlnArgIlePheIleGlyGlnAlaGlnLysPhePheValSerValGlyLeu 320
 DB 940 GCCGTGGATGCGACGAAATATTCMAAGAGCGCAGAAAGTCTTTGATATGTTGGCTTT 999
 QY 321 ProAsnMetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGlyAsnValGln 340

DB 1000 CCTAATATGATCTCAAGGATTTCTGGGAAATTCATGCTAAAGGACCCAGGAAATTTCTAG 1059
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
 DB 1060 AAAGCATGCTGCCATCCACACACTTGGGACCTGGGGGAAAGGGCGACTTACAGATCTCTATG 1119
 QY 361 CysThrLysValThrMetAspPheLeuThrAlaHisIleGlnMetGlyHisIleGln 380
 DB 1120 TGCACAAGGTACATGAGACATTCCTGACAGCTCATCATGAGATGGGCGATATCCAG 1179
 QY 381 TyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsnGlyAlaAsnGlnIlePhe 400
 DB 1180 TATGATATGCAATATGCTGACACACTTTCTGCTAAAGAAATGAGAGCTATATGAAGATTC 1239
 QY 401 HisGlnAlaValGlyGlnIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 DB 1240 CATGAAGCTGTTGGGGAATCATGCTCTCTGACGCGACACTTAACATTTAAATCC 1299
 QY 421 IleGlyLeuLeuSerProAspPheGlnIleAspAsnGlnThrGlnIleAsnPheLeu 440
 DB 1300 ATTTGCTCTTGTACCCGATTTTCAAGAGCAATGAAAGAAATTAACCTTCTGCTC 1359
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnIleTrpArg 460
 DB 1360 AAACAGACACTCAGCATTTGTTGGACTGCGCATTTACTTACTATATGTTAGAGAGTGAGG 1419
 QY 461 TrpMetValPheLysGlyGlnIleProLysAspGlnTrpMetLysLysTrpGlnMet 480
 DB 1420 TGGATGCTCTTAAAGGGGAAATTCCTCAAGAGCCAGTGATGAAAGTGGTGGAGATG 1479
 QY 481 LysArgGlnIleValGlyValAlaGlnProValProHisAspGlnThrTyrCysAspPro 500
 DB 1480 AAGCGAGATATGTTGGGTGGTGGACCTGTGCCCATGATGAAGAACTACTGTGATCCC 1539
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520
 DB 1540 GCATCTCTTCCATGTTTCTGATGATTACCTATTCATTCATTCATTCATTCATTCATTCATTC 1599
 QY 521 TyrGlnPheGlnPheGlnIleValLeuGlnIleValAlaValHisGlnGlyProLeuHis 540
 DB 1600 TACCAATTCACATTCACAAAGACCTTGTCTCAAGCAGCTTAAACATGAAGGCGCTCTGCAC 1659
 QY 541 LysCysAspIleSerAsnSerThrGlnIleGlnIleLysLeuPhe- 555
 DB 1660 AATGTGACATCTCAACTCTACAGAGCTGGACAGAACTTT- GTAAGAAATACCTCA 1718
 QY 555 ----- 555
 DB 1719 AATGTTGAACCTCTCTGATTTACATTCATTCATTCATTCATTCATTCATTCATTCATTC 1778
 QY 555 ----- 555
 DB 1779 ATTTCTTTGTTCTAAAGAAATTTTATGAGCTCAAAATGCTCATTTTACAAACCAA 1838
 QY 555 ----- 555
 DB 1839 CATTTAATTTGGTTCAGACAGAACTTAGACATTAACAATTTGGTGGCCACTCTT 1898
 QY 555 ----- 555
 DB 1899 TTCTCCCTATCAATCACTACAGCCCTCTCTCTGTAATTTGGAAGAAAGAGGCTTATG 1958
 QY 555 ----- 555
 DB 1959 GGTGAATATATCTGTTAATATGATCTTTCTTATCTGACAGAGCAAAATTTAGCCAA 2018
 QY 555 ----- 555
 DB 2019 GTCAAGAGAGAAACCATATGATCATATGATGTAATATATGATGTAATGTAATGTAATGTA 2078
 QY 556 ----- -AsnMetLeuArgLeuGlyLysSerGlnPro 565
 |||||

```

Db 2079 AAGCCCTGAGACCCCTTTTGTGTAGCAATATGCTGAGGCTTGAAATATCAGAACCC 2138
QY 566 TTTTThleuAlaIeuGluAsnValValGlyAlaIeuAsnMetAsnValArgProIleu 585
Db 2139 TGGACCTTGACATGGAATGTTGAGAGC-AAGAAATGAAATGAGGACACGCTC 2197
QY 586 AsnTyrPheGluProIleuPheThrTrpLeuLysAspGluAsnLysAsnSerPheValGly 605
Db 2198 AACTACTTTGAGCCCTTATTTCCTGCGTGAAGACCAAGACAGATCTTTTGCGGA 2257
QY 606 TrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSerLeu 625
Db 2258 TGGAGTACCGACTGAGATCATATGACACCAAGCATCAAGATGAGATTAAGCCTAA 2317
QY 625 sSerAlaIeuGlyAspLysAlaIeuTyrGluTrpAsnAspAsnGluMetTyrLeuPheArg 645
Db 2318 ATCAGCTCTTGAGATTAAGCATATGAAAGCAAGCAATGAAATGACTCTTCGATC 2377
QY 645 rSerValAlaIeuAlaMetArgGlnTyrPheLeuLysValLysAsnGluMetIleuPhe 665
Db 2378 ATCTGTGCATATGCTATGAGCGAGACTTTTAAAAAGTAAAAATCAGATGATCTTT 2437
QY 665 eGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePhe 685
Db 2438 TGGGAGAGAGATGCTCGAGTGCCTAATTGAAACCAAGAAATCTCTTAATTCTTGT 2497
QY 685 lThrAlaProLysAsnValSerAspIleIleProArgTrpGluValGluLysAlaIleAr 705
Db 2498 CACTGACCTAAATATGTTCTGATATCATCTCTGAACTGAAAGTTGAAAGGCAATCG 2557
QY 705 gMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPhe 725
Db 2558 GAGTCCCGAGGCGCTATCATGATGCTTCGCTCGATGACAAACAGCTAGAGTTCT 2617
QY 725 uGlyIleGlnProThrLeuGlyProProAsnGlnProValSerIleTrpLeuIle 745
Db 2618 GGGGATACAGCCCACTTGACCTCTTAACCAAGCCCTGTTCCATATGCTGATGT 2677
QY 745 lPheGlyValAlaMetGlyValIleValGlyIleValIleLeuIlePheThrGly 765
Db 2678 TTTTGGAGTTGTGATGGAGTGTAGTGTGCTGCTGATTCATCTGATTTACGTGGAT 2737
QY 765 eArgAspArgLysLysAsnLysAlaArgSerGlyGluAsnProTyrAlaSerIle 785
Db 2738 CAGAGATCGGAAGAGAAAAATTAACCAAGAGTGAAGAAATTCCTTATGCTCATGCA 2797
QY 785 lPleSerLysGlyGluAsnAsnProGlyPheGlnAsnTrpAspAspValGlnThrSer 805
Db 2798 TATTACCAAGAGAAAAATATCCAGATTCACAAACACTGATGATGTTCAAGACCTCT 2857
QY 805 e 805
Db 2858 T 2858

```

```

; APPLICANT: Matanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; TITLE REFERENCE: P330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 71
; LENGTH: 3732
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-71

Alignment Scores:
Pred. No.: 0 Length: 3732
Score: 4142.00 Matches: 802
Percent Similarity: 85.33% Conservative: 1
Best Local Similarity: 85.23% Mismatches: 2
Query Match: 96.53% Indels: 138
DB: 9 Gaps: 1

US-09-978-385-2 (1-805) x US-10-142-431-71 (1-3732)

QY 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaIleGlnSerThr 20
Db 40 ATGCAACCTCTTCTGCTCTCTCTCTCAGCCCTTCTCTGTAACCTGCTGCTCAGTCCACC 99
QY 21 lIleGluGluGlnAlaIeuThrPheLeuAspLysPheAsnIleGluAlaIeuAspLeuPhe 40
Db 100 ATTGAGAGACAGGCGCAAGACATTTTGGACAGATTTAACCAAGAGCGCAAGACCTGTC 159
QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsnValGln 60
Db 160 TATCAAGTTGACTGCTGCTTCTTGAATTAATTAACCAATATTCTGAAAGAAATGTCCAA 219
QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
Db 220 AACATGAATTAATGCTGGGAGCAAAATGCTGCTCTTTTAAAGGAACGTCACACTTGGC 279
QY 81 GlnMetTyrProLeuGlnIleGlnAsnLeuThrValLysLeuGlnIleGlnAlaLeu 100
Db 280 CAATGTATACCTACACCAAGAAATTCAGATCTCAGATCAAGCTTCAGCTCAGCTCTT 339
QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
Db 340 CAGCAAAATGGGCTCTGAGTCTCTCAGAAAGACAAAGCAAGAGGTTGAACACAAATTC 399
QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnProGlnIle 140
Db 400 AATCAATAGACACACCTCTACGATCTGGAAGTTGTGAACCCAGATTAATCCACAAGAA 459
QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
Db 460 TGCCTATTACTTGAACCAAGCTTGAATGAATATATGCAAAACAGTTTAGCTCAATAGAG 519
QY 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyr 180
Db 520 AGGCTCTGGGCTTGGAAGCTGAGATCTGAGCTCGGCAAGCAAGCTGAGGCTATATAT 579
QY 181 GluGluTrpValValLeuLysAsnGluMetAlaArgAlaAsnIleThrGlyAspTyrGly 200
Db 580 GAAAGATATGCTTGAAGAAATGAGATGGCAAGACCAATCATTTATGAGACTATGGG 639
QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
Db 640 GATTATTGAGAGAGACATGAAATGAATGAGGCTTAATGCTATGCTACACCGCGGC 659
QY 221 GlnLeuIleGluAspValGluIleThrPheGluGluIleLysProLeuTyrGlnIleLeu 240
Db 700 CAGTTGATTGAAGATGAGAAATACATACCTTGAAGAGATTAACCATATATGAAACATCTT 759

```

QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
 |||||
 Db 760 CATGCCATGTGAGGGCAAAAGTTGATGAAAGCCTATCCTCTCATATCAGTCCAAATTGGA 819
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheTyrPheAsnLeuTyrSer 280
 |||||
 Db 820 TGGCTCCCTGGCTCATTTGGCTGGATGTGGGTAGATTTTGGCAAAATCTGTACTCT 879
 QY 281 LeuThrValProPheGlyGlyLeuProAsnIleAspValThrAspAlaMetValAspGln 300
 |||||
 Db 880 TTACAGTTCCCTTTGGACAGAAACCAACATAGATGATTCATGCAATGGTGGACCGAG 939
 QY 301 AlaTyrAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320
 |||||
 Db 940 GCCTGGATGACACAGAAATATTCAGAGAGCGGAGAAATCTCTTGTACTGTGTGCTT 999
 QY 321 ProAsnMetThrGlnGlyThrPheGlyLeuAsnSerMetLeuThrAspProGlyAsnValGln 340
 |||||
 Db 1000 CCTAATATGACTCAAGGATTTGGGAAATTCATGCTTACGACGACCAAGAAATGTTGAG 1059
 QY 341 LysAlaValAlaCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIleLeuMet 360
 |||||
 Db 1060 AAAGCAGTCTGCCATCCACAGCTTGGACCTGGGAGAGGCGACTTCAGATCCCTTAG 1119
 QY 361 CysThrLysValThrMetLysAspPheLeuThrAlaHisGluMetGlyHisIleGln 380
 |||||
 Db 1120 TGCACAAAGGTGACATGACAGCACTTCCTGACAGCTCATGAGATGGGCGATATCCAG 1179
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluPhe 400
 |||||
 Db 1180 TATGATATGGCATATGCTCCACACACTTTCTGCTAAGAAATGGAGCTAATGAAGATTC 1239
 QY 401 HisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 |||||
 Db 1240 CATGAAGCTGTGGGAAATCATGTCACTTCTGACGACACCTGAGGATTTAAATTC 1299
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440
 |||||
 Db 1300 ATTGGCTCTGTCCACCCGATTTTCAAGAACATGAACAGAAATAACTCTCTGCTC 1359
 QY 441 LysGlnAlaLeuThrIleValGlyLysThrLeuProPheThrTyrMetLeuGluLysTyrArg 460
 |||||
 Db 1360 AAACAGACCTCAGCATGTGTGGACTGTGCCATTTACTTACATGTTAGAGAGTGAGAG 1419
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTyrPheGluMet 480
 |||||
 Db 1420 TGGATGTCTTTAAAGGGGAAATTCGCCAAGACCAAGTGTGATGAAAAGTGGGAGATG 1479
 QY 481 LysArgGluIleValGlyValGluProValProHisAspGluThrTyrCysAspPro 500
 |||||
 Db 1480 AAGCGAGAGATAGTTGGGGTGGTGGACCTGTGCCCATGATGAACATACGTGATACCC 1539
 QY 501 AlaSerLeuPheHisValSerAsnAspLysPheIleArgTyrTyrThrArgThrLeu 520
 |||||
 Db 1540 GCATCTCTGTTCCATGTTTGTGATTAATCACTTCACTTCATTCGATTAACACAAAGGACCTT 1599
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuLysGlnAlaIleAlaLysHisGlnGlyProLeuHis 540
 |||||
 Db 1600 TACCAATTCATTCACAGAAAGCACTTGTCTCAAGCAGCTAAACATGAAAGGCCCTGTGAC 1659
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyLysLeuPhe----- 555
 |||||
 Db 1660 AAATGTGACATCTCAAACTCTACAGAGCTGACAGAAACAGTGT- GTAAGAAATACCTCA 1718
 QY 555 ----- 555
 Db 1719 AAATGTGAACCTCTCCTAGATTCAGATTAATCACTTCCATGGCTAGGTTGTATTG 1778
 QY 555 ----- 555
 Db 1779 ATTCTTGTCTAAAGAAATTTTATGGCCTCAAAATGCTCTCATTTACAAACCAA 1838

QY 555 ----- 555
 Db 1839 CATTTAATTTGTGTGACAGACAGAACTAGACCATACAACAATTGGTGGCCACTCTT 1898
 QY 555 ----- 555
 Db 1899 TTCTCCCATATACATACACACCCCTCTCTCTGTGTAATTTGGAAGAAAGCGGTTAG 1958
 QY 555 ----- 555
 Db 1959 GGTGGAATATATCTGTATATGATCTTTCTTATCTGCCAGAACCAATTAGCCAA 2018
 QY 555 ----- 555
 Db 2019 GTCAAGAGAAAGAACCATACATCATAGATGTAAATATATGTACATCTGGAACCCCTCAA 2078
 QY 556 ----- 556
 Db 2079 AAGGCCCTGAACCCCTTTTGTGTGATGCAATATGCTGAGAGCTTGGAAAATACGAACCC 2138
 QY 556 ----- 556
 Db 2139 TGGACCTTAGCATTTGGAATAATGTTGTAGGAGC- AAGAAACATGAATGTAAGGCCACTGCTC 2197
 QY 586 AsnTyrPheGluProLeuPheThrTyrLeuLysAspGlnLysAsnSerPheValGly 605
 |||||
 Db 2198 AACCTACTTGGACCTTATTTTACCTGGCTCAAGACACAGACAGAAATCTTTTGTGGGA 2257
 QY 606 TrpSerThrAspTyrSerProTyrAlaAsp- GlnSerIleLysValArgIleSerLeuLys 625
 |||||
 Db 2258 TGGAGTACCGAGCTGAGTCCATATGACAGACCAAGATCAAAATGAGATAGAGCTTAA 2317
 QY 625 sSerAlaLeuGlyAspLysAlaTyrGluTyrPheLysValLysAsnGluMetIleLeuPhe 645
 |||||
 Db 2318 ATCAGCTCTTGGAGATTAAGCATATGAATGGAACAGACAAATGAATGTACCTGTTCCGATC 2377
 QY 645 rSerValAlaTyrAlaMetArgIleTyrPheLeuLysValLysAsnGluMetIleLeuPhe 665
 |||||
 Db 2378 ATCTGTTGCATATGCTATGAGGCGACTTTTAAAGTAAATAGATGATCTTTT 2437
 QY 665 eGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePheVal 685
 |||||
 Db 2438 TGGGAGAGGAGATGTGGAGTGGCTAATTGGAACCAAGATCTCCTTATTTCTTGT 2497
 QY 685 lThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAlaIleArg 705
 |||||
 Db 2498 CACTGCACCTTAAATAATGTGTGATATCATTTCTTAAGAACTGAAGTTGCAAAAGGCCATCAG 2557
 QY 705 gMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLe 725
 |||||
 Db 2558 GATGTCCCGAGGCGGTATCAATGATGCTTTCCTGCTGAATGACACAGCTTAGAGTTTCT 2617
 QY 725 uGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerIleThrPheLeuLeuVal 745
 |||||
 Db 2618 GGGGATACAGGCAACACTGTGACCTCTTAACCAAGCCCCCGTTTCCATATGGCTGATGT 2677
 QY 745 lPheGlyValValMetGlyValIleValValGlyIleValIleLeuIlePheThrGlyI 765
 |||||
 Db 2678 TTTTGGAGTTGTGATGGAGAGTGAATGATGTGGCATTTGTCATCTGCTTCACTGGGAT 2737
 QY 765 eArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnProTyrAlaSerIleAs 785
 |||||
 Db 2738 CAGAGATCGGAGAGAAAGAAATTAAGCAGAAAGTGGAGAAATCTTATGCTCCATCGA 2797
 QY 785 pIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPhe 805
 |||||
 Db 2798 TATTAGCAAGAGAAATAATATCCAGAGATTCAAAACACTGATGATGTTCCAGACCTCCTT 2857
 QY 805 e 805
 Db 2858 T 2858
 RESULT 13

```

US-10-143-114-71
; Sequence 71, Application US/10143114
; Publication No. US20030036180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33303R1C211
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 71
; LENGTH: 3732
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-143-114-71

Alignment Scores:
Pred. No.: 0 Length: 3732
Score: 4142.00 Matches: 802
Percent Similarity: 85.338 Conservative: 1
Best Local Similarity: 85.238 Mismatches: 2
Query Match: 96.538 Indels: 138
Gaps: 1

US-09-978-385-2 (1-805) x US-10-143-114-71 (1-3732)
QY 1 MetSerSerSerSerTrrPleuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
DB 40 ATGTCAAGCTCTTCTGCGCTCTTCACAGCTTGTGTAAGTCTGCTCAGTCCAC 99
QY 21 ILGLIUGLUGLAlaLysThrPheLeuAspLysPheAsnHisGlnLysGlnLysPhe 40
DB 100 ATGAGAGAACAGGCCAAGACATTTTGGACAAAGTTTACCCAGCAAGCCGAGACCTGTT 159
QY 41 TyrGlnSerSerLeuAlaSerTrrPasnTyrAsnThrAsnIleThrGlnLysValGln 60
DB 160 TATCAAGAAGTCACTTCTCTTCTGCAATTAACACCAATATTACTGAACAGATGTC 219
QY 61 AsnMetAsnAsnAlaGlyAspLysTrrPserAlaPheLeuLysGlnGlnSerThrLeuAla 80
DB 220 AACATGAATTAATGCTGGGCAAAATGCTGCTTTTAAAGAAACAGTCCACACTTGC 279
QY 81 GlnMetTrrProLeuGlnLysIleGlnAsnLeuThrValLysLeuGlnLysAlaLeu 100
DB 280 CAATATATCCACTACAAAGAAATTCAGATCTCACAGTCAAGCTTACGTCAGGCTCT 339
QY 101 GlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThrIleLeu 120
DB 340 CACCAAAATAGGCTTCACTGCTCTCAGAAAGCAAGCAAGCAAGGTTGAACAAATTTCT 399
QY 121 AsnThrMetSerThrIleLeuSerThrGlyLysValCysAsnProAspAsnProGlnGln 140
DB 400 AATATCAATAGCACCATCTACAGTACTGAAAAAGTTTGTAAACCCAGATATCCACAAAGA 459
QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTrrAsnGln 160

```

```

DB 460 TECTATTACTTGAAACAGGTTTGATGAATTAATGCAAAACAGTTTACTACATAGCAG 519
QY 161 ArgLeuTrrPalaTrrPLeuSerTrrPArgSerGlnValGlyLysGlnLeuArgProLeuTyr 180
DB 520 AGGCTCTGGGCTTGGGAAAGCTGGAGATTCAGGCTGGGCAAGCAGCTGAGCCATTATAT 579
QY 181 GlnGluTrrValValLeuLysAsnGlnLysValAlaArgAlaAsnHisTyrGlnAspTrrGly 200
DB 580 GAAAGATATGCTGTTTAAAAAATGATGGCAGAGCAAAATCTTATGAGGACTATGG 639
QY 201 AspTrrTrrPArgGlyAspTrrGlnValAsnGlyValAspGlyTrrAspTrrSerArgGly 220
DB 640 GATTTATGGAGAGGAGACTATGAAGTAATAGGGAGATGCTATGACTACAGCCGGGCG 699
QY 221 GlnLeuIleGlnAspValGlnHisThrPheGlnGlnIleLysProLeuTyrGlnHisLeu 240
DB 700 CAGTTGATTTGAAGATGGTGAACATACCTTGAAGAGATTTAAACCATATATGAACATCTT 759
QY 241 HisAlaTrrValArgAlaLysLeuMetAsnAlaTrrProSerTrrLysSerProIleGly 260
DB 760 CATGCTTATGGAGGCAAAAGTTGATGAATGCTTATCTTATATCAGTCCATTTGCA 819
QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrrPArgPheTrrPThrAsnLeuTrrSer 280
DB 820 TGCCCTCCCTGCTCATTTGCTTGTGATGTGGGGTAGATTTTGGACAAATCTGACTCT 879
QY 281 LeuThrValProPheGlnGlnLysProAsnIleAspValTrrAspAlaMetValAspGln 300
DB 880 TTGACAGTTCCTCTTGGGACAGAAACCAACATAGATGTTACTGATGCAATGGTGGACAC 939
QY 301 AlaTrrAspAlaGlnArgIlePheLysGlnLysGlnLysPhePheValSerValGlyLeu 320
DB 940 GCCTGGGATGCACAGAGAAATTTCAAGAGGCCAGGAAGTCTTGTATCTGTTGGCTT 999
QY 321 ProAsnMetThrGlnGlnLysPheTrrPLeuAsnSerMetLeuThrAspProGlyAsnValGln 340
DB 1000 CCTAATATGACTCAAGGATTTGCGAAATTCATGCTTAACAGGACCCAGCAAAAGTTCAG 1059
QY 341 LysAlaValCysHisPrrThrAlaTrrAspLeuLysGlyAspPheArgIleLeuMet 360
DB 1060 AAAGCAGTGTGCATCCACACAGCTTGGGACCTGGGGAAGGCGCACTCAGATCTTATG 1119
QY 361 CysThrLysValThrMetAspPheLeuThrAlaHisHisGlnMetGlyHisIleGln 380
DB 1120 TGCACAAAAGGTCACATGAGACATCTCTGACACTCATGATGAGTGGGCAATATCCAG 1179
QY 381 TyrAspMetAlaTrrValAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnLysPhe 400
DB 1180 TATGATATGGCATATGCTGCACAAACCTTTTCTGCTAAGAAATGAGCTAATGAAGATTC 1239
QY 401 HisGlnAlaValGlyGlnIleMetSerLeuSerAlaIleThrProLysHisLeuLysSer 420
DB 1240 CATGAGAGCTGTGGGGAATCATGTCACCTTCTGCACGCCACACTTAAGCATTTAAATCC 1299
QY 421 IleGlyLeuLeuSerProAspPheGlnGlnAspAsnGlnThrGlnIleAsnPheLeuLeu 440
DB 1300 ATTGCTCTTCTGTCACCCGATTTTCAAGAAACATGAAGAAACGAAATAAATCTTCTC 1359
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheTrrTrrMetLeuGlnLysTrrArg 460
DB 1360 AAACAGACACTCAGATTTGGGACCTGTCATTTACTTATCATGTTAGAGAACTGGAGG 1419
QY 461 TrrMetValPheLysGlyGlnIleProLysAspGlnTrrPheLysLysTrrPrrGlnMet 480
DB 1420 TGGATGGCTTTAAAGGGGAATTCACCAAGACAGAGGATGAAGAAAGTGTTGGAGATG 1479
QY 481 LysArgGlnIleValGlyValAlaGlnProValProHisAspGlnTrrTrrCysAspPro 500
DB 1480 AAGGAGAGATAGTTGGGGTGGTGAACCTGTGCCCATATGAAACATCTGAGACCCC 1539
QY 501 AlaSerLeuPheHisValSerAsnAspTrrSerPheIleArgTrrTrrArgThrLeu 520

```


QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnInserThrLeuAla 80
 |||||
 Db 220 AACATGATATATGCTGGGAGCAAAATGCTCTCTTTTAAAGGAACAGTCCACTTGCC 279
 QY 81 GlnMetYrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuAlaLeu 100
 |||||
 Db 280 CAAATGTCACACTACAGAAATTCAGAAATCTCAGCAAGCTTCAGCTCAGCTCTT 339
 QY 101 GlnGlnAsnGlySerSerValLeuSerGlnLysPylSerLysAspGluAspThrIleLeu 120
 |||||
 Db 340 CAGCAAAATGGCTCTTCAGTCTCTCAGAAAGCAAGCAAAAGCGTTGAACAAATCTTA 399
 QY 121 AsnThrMetSerThrIleLeuSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
 |||||
 Db 400 AATAAAGAGAGACCATCTACAGTCTGAGAAAGTTGTAAACCCAGAAATCCCAAGAA 459
 QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
 |||||
 Db 460 TGCCTTATTAAGTGAACACAGGTTTGAATGAATGAACAGTTTGAAGTGAACATGAG 519
 QY 161 ArgLeuThrAlaTrpGlnSerTrpArgSerGluValGlyLysGlnLeuArgProLeuThr 180
 |||||
 Db 520 AGGCTGTGGGCTGGGAAAGCTGGAGATCTGAGGTCGCAAGAGCTGAGGCTCATTTAT 579
 QY 181 GlnGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200
 |||||
 Db 580 GAAAGATATGTCGCTTAAAAATAGATGAGCAAGCAAAATCATTAAGAGACTATGGG 639
 QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGluValAspGlyTyrAspTyrSerArgGly 220
 |||||
 Db 640 GATTTATGGAGAGAGACTATGAAGTAATGAGGAGTATGCTATGACTACAGCCGCGC 699
 QY 221 GlnLeuIleGlnAspValGluHisThrPheGlnGluIleLysProLeuTyrGlnHisLeu 240
 |||||
 Db 700 CAGTGTATGAGATGTGGAACATACCTTTGAGAGATTAACCATTAATGAACATCTT 759
 QY 241 HisAlaTyrValAlaGlnLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
 |||||
 Db 760 CATGCTATGTCAGGCAAGTGTGATGATGATGCTCTCTCTATATCAGTCCCAATTGGA 819
 QY 261 CysLeuProAlaHisLeuLeuGlnLysPylSerTrpGlyArgPheThrAsnLeuTyrSer 280
 |||||
 Db 820 TGCCTCCCTGCTCATTTCTGCTGATGATGCTGAGTATTTGGCAAAATCTGACTCT 879
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
 |||||
 Db 880 TTGACAGTTCCTTTGGACAGAAACCAACATAGATGTACTGATGATGATGATGATGATG 939
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerValGlyLeu 320
 |||||
 Db 940 GCGTGGATGACACAGAAATATTCAGAGAGCGCGAGAAAGTCTTTGTATCTGCTGGCTT 999
 QY 321 ProAsnMetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGlyAsnValGln 340
 |||||
 Db 1000 CTTATATGACTCAAGGATTTCTGGGAAATTCATGATGCTAAAGGAGCCGAGAAATGTTCCAG 1059
 QY 341 LysAlaValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
 |||||
 Db 1060 AAAGAGCTGCTGCATCCACAGCTGGAGCTGGGAAAGGCGACTTGAAGATCTTATG 1119
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisLysGlnMetGlyHisIleGln 380
 |||||
 Db 1120 TGCACAAAGAGTACAAATGAGACACTTCTGACAGCTCATGAGATGGGCAATATCCAG 1179
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuArgAsnGlyAlaAsnGlnGlyPhe 400
 |||||
 Db 1180 TATGATATGGCATATGTCGACAAACCTTTCTGCTAAGAAATGAGCTAATGAGGATTC 1239
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
 |||||
 Db 1240 CATGAAGCTGTGGGAAATCATGTCACTTCTGACAGCCACACCTTAACATTTAAATCC 1299

QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440
 |||||
 Db 1300 ATTGCTCTCTGTCACCCGATTTTCAAGAAAGCAATGAAGAAACAGAAATTAATCTCTGCTC 1359
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460
 |||||
 Db 1360 AAACAGACATCCAGATGTGGAGCTGCGCAATTTACTTACATGTGTAAAGAAAGTGAAG 1419
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysTrpTrpGlnMet 480
 |||||
 Db 1420 TGGATGCTCTTTAAAGGGAATTTCCCAAGACCATGATGAAAAAGTGGGAGATG 1479
 QY 481 LysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCysAspPro 500
 |||||
 Db 1480 AAGCAGAGATGATGGGGGAGTGAACCTGTGCCCATGATGAAGAAATCATCTGTGACCC 1539
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520
 |||||
 Db 1540 GCATCTCTGTCATGTTCTGTGATGATCTCATTTCTGATATTACAAAGAGACCTT 1599
 QY 521 TyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHis 540
 |||||
 Db 1600 TACCAATTCAGATTTCAAGAGACACTTTGTCAAGCAGCTTAACATGAAGCCCTCTGAC 1659
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 555
 |||||
 Db 1660 AAATGTACATCTCAAACTCTACAGAAAGCTGACAGAAACTGTT-CTAAGAAATACCTCA 1718
 QY 555 ----- 555
 Db 1719 AAATGTTGAACCTCTCTAGTATTCAGTATTCATCTTCATGCTAGCTTGTATTTG 1778
 |||||
 QY 555 ----- 555
 Db 1779 ATTTCTTTGTTCAAAAAGAAATTTTATGCGCTCAAAATGTCCTCATTTACAACCAAA 1838
 |||||
 QY 555 ----- 555
 Db 1839 CATTTATTTGTGTGTCAGACAGAACCTAGACCATCAACAATTTGGTGGCCACCTT 1898
 |||||
 QY 555 ----- 555
 Db 1899 TTCTCCCTATCATTAACACAGCCCTCTCTCTGTAATTTGAAGAAAGCGGTTAG 1958
 |||||
 QY 555 ----- 555
 Db 1959 GGTGGAATATATGTTAATATGATCTTTTCTTATCTGCCAGAAAGCAAAATTTAGCCAA 2018
 |||||
 QY 555 ----- 555
 Db 2019 GTCAAGAGAAAGAACATGATCATGATGTAATATATGTCATCTGGAACCCCTCAA 2078
 |||||
 QY 556 ----- 556
 Db 2079 AAGGCCCTGAACCCCTTTTGTGTAGCAAAATCTGAGGCTTGGAAATCAGAAACC 2138
 |||||
 QY 566 TrpThrLeuAlaLeuGlnAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeu 585
 |||||
 Db 2139 TGGACCTTACATTTGAAATGTTGTGAGAC- AAGACATGATATTAAGGCCACAGCTC 2197
 |||||
 QY 586 AsnTyrPheGlnProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGly 605
 |||||
 Db 2198 AACACTTTAGGCCCTTATTTACCTGGCTGAAGAGACAGAACAAATCTTTTGGGGA 2257
 |||||
 QY 606 TrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSerLeuLys 625
 |||||
 Db 2258 TGGAGTACCACATGAGTCCATATGACAGACCCCAACATCAAGTGAAGATTAAGCTTAA 2317
 |||||
 QY 625 sSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGlnMetTyrLeuPheArgse 645
 |||||
 Db 2318 ATCAGCTCTTGAGATTAACATATGAATGAAGCAATGAATGATTTACTGTTCCGATC 2377
 |||||
 QY 645 rSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetIleLeuPhe 665

```

Db 2378 ATCTGTGATATGCTATGAGCAGACTTTTAAAGAAAAATCAGATCTTTT 2437
Qy 665 eelglugluaspvalargvalalalsnleuysproargileserpheasnphpheva 685
Db 2438 TGGGGAGGAGATGTGCGAGTGGCTAAATTGAAACCAAGAAATCTCTTAATTCTTTGT 2497
Qy 685 lthrlap:polysasnvalseraspilieleproarghrgrlvalglulysalallear 705
Db 2498 CACTGCACCTAAAAATGTGTCTGATATATCTTCTTAAGAACTGAAGTTGAAGGCCATCAG 2557
Qy 705 gnetserargserargileasnspalapheargleuasnspasnserleugluphele 725
Db 2558 GATGTCCGGAGCCGATCAATGATGCTTCCGTCTGATGACAAACACCTAGATTTCT 2617
Qy 725 ugllytlegrnprothrleuglyproproasnleuprovalseriletrpleullea 745
Db 2618 GGGGATACAGCCACACTTGGACCTTACACAGCCCTGTTCCATATAGCTGATTTGT 2677
Qy 745 lphelglvalvalmetglvalillevalvalgllyllevalilleuilephethrgrl 765
Db 2678 TTTTGATTTGATGGAGAGATGATGTGTTGGCATTTGTCATCTTGATCTTCACTGGGAT 2737
Qy 765 eargaspargllysllysasnlysalaargserglgluasnprotyralaserileas 785
Db 2738 CAGACATCGAAGAAAGAAATAAAGCAAGAAGTGAGAAAAATCCTTATGCTTCATCGA 2797
Qy 785 pleserlysglgluasnasnprogllyphelglasnthrspaspyalglthrserph 805
Db 2798 TATTTCAGAAAGAGAAATAATCAAGATTCACAAACACTGATGATGTCAGACCTCTT 2857
Qy 805 e 805
Db 2858 T 2858

```

RESULT 15
US-09-969-384-2
; Sequence 2, Application US/09969384
; Publication No. US20020192749A1
; GENERAL INFORMATION:
; APPLICANT: Moore, et al.
; TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PTO55PI
; CURRENT APPLICATION NUMBER: US/09/969,384
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: PCT/US01/10542
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/236,384
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/194,118
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1707)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2702)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2749)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2757)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2788)
; OTHER INFORMATION: n equals a,t,g, or c

```

; NAME/KEY: SITE
; LOCATION: (2789)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2819)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2835)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2856)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-969-384-2

Alignment Scores:
Pred. No.: 0
Score: 4061.00
Percent Similarity: 99.35%
Best Local Similarity: 99.22%
Query Match: 94.64%
DB: 9
Gaps: 0
Length: 2920
Matches: 763
Conservative: 1
Mismatch: 4
Indels: 1

US-09-978-385-2 (1-805) x US-09-969-384-2 (1-2920)
Qy 3 SerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThrIleGlu 22
Db 35 AGCTCTCCCTGGCTCTTCTCACCTGTTGGCTGTAACGTGCTGATCCACCATTTGAG 94
Qy 23 GluGlnAlaLysThrPheLeu-AspLysPheAsnHISGluAlaGluAspLeuPheTrgl 42
Db 95 GAACAGGCCAAGACATTTTGGGACCAAGTTTAAACCGAAGCCGAAGACCTGTTCTATCA 154
Qy 42 nSerSerLeuAlaSerTrpAsnTrpAsnThrAsnIleThrIleGluGluAsnValGlnAsn 62
Db 155 AAGTTCACCTGCTTCTTGGATATTAACCAATATTACTGAAGAGATGTCACAAACAT 214
Qy 62 lAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAlaGln 82
Db 215 GAATTAATGCTGGGGACAAATAGCTCTCTTTTAAAGGAACAGTCCACACTTGCCAAAT 274
Qy 82 tTryProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuAlaLeuGln 102
Db 275 GATTCACACTACAGAAATTCAGAAATCTCAGATCTCAGATCAAGCTTCAGAGCTTCA 334
Qy 102 nAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeuAsnTh 122
Db 335 AAATGGGCTTTCAGTCTCTCAGAAAGCAAGCAAAACGGTGAACACATAATCTAAATAC 394
Qy 122 rMetSerThrIleTrySerThrGlyLysValCysAsnProAspAsnProGlnGluCysLe 142
Db 395 AATGACACACATCTACAGTACTGAAAAAGTTGTAAACCCAGATTAATCCACAGAAATGCTT 454
Qy 142 uLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTryAsnGluArgLe 162
Db 455 ATTACTTGAACACAGGTTTGAATGAATAATGCAAAACAGTTAGACTCAATGAGAGGCT 514
Qy 162 uTrpAlaTrpGluSerTrpArgSerGluValAlaGlySglnLeuArgProLeuTryGlu 182
Db 515 CTGGGCTTGGGAAAGCTGGAGATCTGAGGTCGGCAAGCAGCTGAGCCATTAATGAAGA 574
Qy 182 uTryValValLeuLysAsnGluMetAlaArgAlaAsnHISTryGluAspTryGlyAspTry 202
Db 575 GATGTGTCCTTGAATAATGAGATGGCAAGACCAATCATTTATGAGACTATGGGATTA 634
Qy 202 rTrpArgGlyAspTryGluValAsnGlyValAspGlyTryAspTrySerArgGlyGlnLe 222
Db 635 TTGGAGAGGAGACTATGAAGTAATGGGTAGATGCTATGACTACACCCGGCCAGACT 694
Qy 222 uLleGluAspValGluHISThrPheGluGluIleLysProLeuLysGluHISAla 242
Db 695 GATTCAGATGTGGAACATACCTTTGAAGAGATTTAAACCATTAATGAACATCTTCATGC 754
Qy 242 aTryValAlaGluAlaLysLeuMetAsnAlaTryProSerTryIleSerProIleGlyCysLe 262

```

```

|||||
Db 755 CTATGTAGGCGCAAGTTGATGATGCTTCTCTTCTATATGATCCATTTGGATGCTT 814
QY 262 uPrOAlAhISleuLeuGlYAspMetTrpGIYArpPheTrpThrAsnLeuTySerLeuTh 282
Db 815 CCCGCGCATTTGCTTGGTATGATGTTGGAGATTTTGGACAATAATGTACMSTTTGAC 874
QY 282 rValProPheGlyGlnYsProAsnIleAspValThrAspAlaMetValAspGlnAlaTr 302
Db 875 AGTTCCCTTTGGACGAAGCAACCAATAGATTGATGATGCAATGGTGGACCAAGRCCTG 934
QY 302 pAspAlaGlnArGlyIlePheLysGlnAlaGluLysPhePheValSerValGlyLeuProAs 322
Db 935 GGATGCAAGAGATATTCAAGAGGCCGAGAACTTTCTTGTATCTTGGTCTCTCTAA 994
QY 322 nMetTrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGlnLysAl 342
Db 995 TATGACTCAAGATTCGTGGAAATTCATGCTAACGACCCAGAAATGTTCCAGAAAGC 1054
QY 342 aValCySHISProThrAlaTrpAspLeuGlyLysGlyAspPheArGlyIleuMetCysTh 362
Db 1055 AGTGGCCATCCCAAGCTTGGAGCTGGGGAAGGGGACTTCAGATCTTATGTGCAC 1114
QY 362 rLysValThrMetAspAspPheLeuThrAlaHISISGluMetGlyHISISleGlnTyAs 382
Db 1115 AAAGTGCATATGACGACCTTCGACAGCTCATGAGATGGGCAATATCCAGTATGA 1174
QY 382 pMetAlaTyAlaAlaGlnProPheLeuLeuArGAsnGlyAlaAsnGluGlyPheHISG 402
Db 1175 TATGCAATATGCTGCACAACTTTCTGCTAAGAAATGAGCAATGAAAGATTCATGA 1234
QY 402 uAlaValGlyGlnIleMetSerLeuSerAlaAlaThrProLysHISLeuLysSerIleG 422
Db 1235 AGCTGTTGGGAATCATCATCTTCTGACGCCACACTAGCAATTTAAATCATTCG 1294
QY 422 yLeuLeuSerProAspPheGlnGluAspAsnGluThrGlnIleAsnPheLeuLeuLysG 442
Db 1295 TCTTGTGACCCGATTTTCAAGAAAGACATGAAGAAATAAATCTGCTCAACA 1354
QY 442 nAlaLeuThrIleValGlyThrLeuProPheThrTyMetLeuGlnLysTrpArgTrpMe 462
Db 1355 AGCACTCAAGATTTGGGACTGCTGCATTTACTTACATGTTAGAGAGTGGAGTGAT 1414
QY 462 tValPheLysGlyGlnIleProLysAspGlnTrpMetLysTrpTrpGluMetLysAr 482
Db 1415 GGTCTTTAAAGGGAATTTCCAAAGACAGTGGATGAAAAAGTGTGGAGATGAAGCG 1474
QY 482 gGluIleValGlyValGluProValProHISAspGluThrTyCysAspProAlaSe 502
Db 1475 AGAGATAGTGGGTGGTGGAACTGTGCCCATGATGAAACATGCTGTGACCCCGCATC 1534
QY 502 rLeuPheHISValSerAsnAspTySerPheIleArGlyTyTrpThrArgTrpIleuTyG 522
Db 1535 TCTGTTCCATGTTTCTAATGATTACTCATTTCTGATATTACACAAAGGACCTTTACCA 1594
QY 522 nPheGlnPheGlnGluAlaLeuGlnAlaAlaLysHISGluGlyProLeuHISLysCy 542
Db 1595 ATTCCAGTTTCAAGAGCACTTGTCAAGCAAGCTAAACATGAAGCCCTCGACAAATG 1654
QY 542 sAspIleSerAsnSerThrGlnAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGly 562
Db 1655 TGACATCTCAAACTACAGAGAGTGGACAGAAACTGTTCAATATGCGAGAGNTGGAAA 1714
QY 562 sSerGluProTrpThrIleuAlaLeuGlnAsnValAlaGlyAlaLysAsnMetAsnValAr 582
Db 1715 ATTCAGAACCTTGACCTAGCATGGAATGTTGTAGAGCAAAACAATGATGTAAG 1774
QY 582 gProLeuLeuAsnTyPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSe 602
Db 1775 GCCACGCTCAACTACTTGGACCCCTATTTACCTGGGTGAAAGCAAGCAAGAAATTC 1834
QY 602 rPheValGlyTrpSerThrAspTrpSerProTyAlaAspGlnSerIleLysValArgI 622
|||||

```

```

Db 1835 TTTTGGGATGAGTACGACTGAGTCCATATGACAGCAACCAATCAAAAGTGAGAT 1894
QY 622 eSerLeuLysSerAlaLeuGlyAspLysAlaTyGlnTrpPAsnAspAsnGluMetTyLe 642
Db 1895 AAGCCTAAATATGAGCTTTGGAGATTAAGCATATGAAATGAACCAATGAATGACTT 1954
QY 642 uPheArGSerSerValAlaTyAlaMetArGlnTrpPheLeuLysValLysAsnGlnMe 662
Db 1955 GTTCCGATCATCTGTTGCAATATGATGAGCAGTACTTTTAAAGTAAAAATCAGAT 2014
QY 662 tIleLeuPheGlyGlnGluAspValArgValAlaAsnLeuLysProArGlyIleSerPheAs 682
Db 2015 GATTCCTTTTGGGAGAGAGATGTGCGAGTGTATTTGAACCAAGAAATCTCTTTAA 2074
QY 682 nPhePheValThrAlaProLysAsnValSerAspIleIleProArGTrpGlnGluValG 702
Db 2075 TTTCTTTGCTACTGCACTTAAATATGCTGTATATCATCTTCAACTGAAGTGAATA 2134
QY 702 sAlaIleArGMetSerArGSerArGTrIleAsnAspAlaPheArGLeuAsnAspAsnSerLe 722
Db 2135 GGCCATCAGAGATGTCGCCGAGCCGTATCAATGATGCTTCCGTCGAATGACGACAGCT 2194
QY 722 uGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerIleTr 742
Db 2195 AGAGTTTCTGGGATACACCAACACTTGACCTCTTAACAGCCCTGTTCCATATG 2254
QY 742 pIleuIleValPheGlyValAlaMetGlyValIleValAlaGlyIleValIleLeuIlePh 762
Db 2255 GCTGATTTGTTTGGAGTTGTATGAGGAGTGAATGAGTGGCATTTGCTATCTGATCT 2314
QY 762 eThrGlyIleArGAspArgLysLys 770
Db 2315 CACTGGATCAGAGATCGAAGAG 2339

RESULT 16
US-09-969-384-12
; Sequence 12, Application US/09969384
; Publication No. US20020192749A1
; GENERAL INFORMATION:
; APPLICANT: Moore, et al.
; TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT055P1
; CURRENT APPLICATION NUMBER: US/09/969,384
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: PCT/US01/10542
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/236,384
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/194,118
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 2911
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1705)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2847)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-969-384-12

Alignment Scores:
Pred. No.: 0
Score: 4013.00
Percent Similarity: 99.35%
Best Local Similarity: 99.22%
Query Match: 93.52%
Gaps: 0
Length: 2911
Matches: 763
Conservative: 1
Mismatch: 4
Indels: 3
Gaps: 0

```

US-09-978-385-2 (1-805) x US-09-969-384-12 (1-2911)

QY 3 SerSerSerThrPheLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThrIleGlu 22
 |||||
 Db 35 AGCTCTTCGGCTCCCTTCACGCTTGTGCTGAACGCTGCTCAGTCACACATTTAG 94
 QY 23 GluGlnAlaIAspThrPheLeu-AspLysPheAsnHisGluAlaGluAspLeuPheArgI 42
 |||||
 Db 95 GAACAGGCCCAAGACATTTTGGGACACAAGTTTAAACCACGAGGCCGAGACCTGTTTATCA 154
 QY 42 nSerSerLeuAlaSerTrpAsnTrpAsnThrAsnIleThrGluGluAsnValGlnAsnMe 62
 |||||
 Db 155 AAGTTCACTTGTCTTGGAAATTATTAACACCATATTACTGAGAGAAATGTCCAAAACAT 214
 QY 62 LAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAlaGlnMe 82
 |||||
 Db 215 GAATTAATGCTGGGACAAATGCTCCTTTTAAAGAACACTCCACCTGCCCCAAAT 274
 QY 82 tTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaIleGlnGln 102
 |||||
 Db 275 GTATCCACTACAGAAATTCAGAAATTCACAGTCAAGCTTCAGCTCAGGCTCTTACGCA 334
 QY 102 nAsnGlySerSerValLeuSerLysLysLysSerLysArgLeuAsnThrIleLeuAsnTh 122
 |||||
 Db 335 AAATGGGTCTTCAGTCTCTCAGAACAGACAAAGCAACGTTGAAACAAATTCCTAAATAC 394
 QY 122 rMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGluCysLe 142
 |||||
 Db 395 AATGACACACCATTCACAGTACAGTACGAAATTTGTATACCCAGATTAATCCAAAGAAATGCT 454
 QY 142 uLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGluArgLe 162
 |||||
 Db 455 ATTACTTGAAACCGATTGAATGAATATATGCGCAACAGCTTAAAGCTACAAATGAGAGGCT 514
 QY 162 uTrpAlaTrpGlySerTrpArgSerLysValGlyLysGlnLeuArgProLeuTyrGluGln 182
 |||||
 Db 515 CTGGGCTTGGGAAAGCTGAGATCTGAGGCTGGCAGAGCTGAGCCCTTTATATGAGAA 574
 QY 182 uTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGlyAspTyr 202
 |||||
 Db 575 GTATGTGGTCTTGAATAATGAGATGCGACAGACCAATCTTTATGAGAGCTATGGGATTA 634
 QY 202 rTrpAlaGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGlyGlnLe 222
 |||||
 Db 635 TTGGAGAGGAGACTATGAGATGAATGGGCTAGATGGCTATGACTACAGCCGGCCAGT 694
 QY 222 uIleGluAspValGluHisThrPheGlnGluIleLysProLeuTyrGlnHisLeuHisAl 242
 |||||
 Db 695 GATTCAGATGTGGACATACCTTTGAAAGAGATTAACCCATTAATGAAACATCTTCATGC 754
 QY 242 aTyrValAlaArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGlyCysLe 262
 |||||
 Db 755 CTATGTGAGGCCAAAGTTGATGATGCCATCTCCATATCACTCCAAATTTGGATGCT 814
 QY 262 uProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTyrSerLeuTh 282
 |||||
 Db 815 CCCTGCTCATTTGCTTGGATATGTGGGCTAGATTTGGACAATCTGTACTCTTGGAC 874
 QY 282 rValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGlnAlaTr 302
 |||||
 Db 875 ACTTCCCTTTGGACAAACCAACATAGATGTACTGATGCAATGGTGGACAGGCTGTG 934
 QY 302 pAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeuProAs 322
 |||||
 Db 935 GGATCCACAGAGAAATATTCAGAGAGCGGAGAGTCTT-GRATGTGTGGCTTCTCTTA 993
 QY 322 nMetThrGlnGlyPheTrpGlyAsnSerMetLeuThrAspProGlyAsnValGlnLysAl 342
 |||||
 Db 994 TATGACTCAAGGATTTCTGGGAAATTCATGCTTAACGAGACCAGGAAATGTTCAGAAAGC 1053
 QY 342 aValAlaLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMetCysTh 362
 |||||
 Db 1054 ACTTGCATCCCAAGCTTGGACCTGGGAGGGCGACTTCAGAGATCCCTTATGTGTGAC 1113

QY 362 rLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlnHisIleGlnTyrAs 382
 |||||
 Db 1114 AAAGTGACCAATGACGACCTTCAGACGCTCATATGAGATGGGCAATATCCAGTATAG 1173
 QY 382 pMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyPheHisGln 402
 |||||
 Db 1174 TATGGCATATGCTGCACAACCTTTCTGCTAAAGAAATGAGCTAATGAAAGATTCATAGA 1233
 QY 402 uAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSerIleG 422
 |||||
 Db 1234 AGCTGTGGGCAATCATCTCTCTTCTGCAGCCACACCTTAAGCATTTTAAATTCATTTGG 1293
 QY 422 yLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLysGln 442
 |||||
 Db 1294 TCTTCTGCACCCGATTTTCAAGAAAGACATGAACAAATTAACCTTCGTGCTCAAAACA 1353
 QY 442 nAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTrpArgTyrPme 462
 |||||
 Db 1354 AGCACTCACGATGTTGGGAGCTGTGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1413
 QY 462 tValPheLysGlyGluIleProLysAspGlnTrpMetLysTrpTrpGluMetLysArg 482
 |||||
 Db 1414 GGTCTTTAAAGGGGAAATTCGCCAAAGACACAGTGAATAAGTGTGGAGATGAAGCG 1473
 QY 482 gGluIleValGlyValGluProValProHisAspGluThrTyrCysAspProAlaSe 502
 |||||
 Db 1474 AGAGATAGTGGGGTGGGAACTGTGCCCATGATGAAGAAACATATGTCGACCCGATC 1533
 QY 502 rLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrpThrArgThrLeuTyrGln 522
 |||||
 Db 1534 TCTGTTCATGTTCTTAATGATTAATCTCATTTATTCGATTAATCAAGAGACCTTTACCA 1593
 QY 522 nPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHisLysCy 542
 |||||
 Db 1594 ATTCAGATTCAGAGACACTTGTGCAAGCAGTAAACATGAAGCCCTGCGCAAAATG 1653
 QY 542 sAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLy 562
 |||||
 Db 1654 TGACATCTGC-AAGCTACAGAGAGCTGGACAGAAACCTGTCAATATGCTGAGGNTGGAA 1712
 QY 562 sSerGlnProTrpThrLeuAlaLeuGluAsnValValGlyValAlaLysAsnMetAsnVal 582
 |||||
 Db 1713 ATCAGAAACCCGAGACCTGACATTTGGAATGTTGTGAGACCAAGACATGAATGTAAG 1772
 QY 582 gProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSe 602
 |||||
 Db 1773 GCCACTGCTCAACTATTTGAGCCCTTATTTACTGCTGCGTGAAGACCAAGAAATTC 1832
 QY 602 rPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysValArgIle 622
 |||||
 Db 1833 TTTTGTGGAGATGAGTACCGAGCTGAGTCCATATGACAGACCAAGCTCAAAATGAGGAT 1892
 QY 622 eSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLe 642
 |||||
 Db 1893 AAGGCTAAATACACTCTTGGAGATTAAGCATATGATGAGACAGACATGAATGTAACCT 1952
 QY 642 uPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMe 662
 |||||
 Db 1953 GTTCCGATCATCTGTTGCATATGCTATGAGGAGAGTACTTTTAAAGTAAATAAATCAGAT 2012
 QY 662 tIleLeuPheGlyGlnGluAspValArgValAlaAsnLeuLysProArgIleSerPheAs 682
 |||||
 Db 2013 GATTCCTTTTGGGAGGAGATGAGGAGGCTTAATTTGAACCAAGAAATTCCTTTAA 2072
 QY 682 nPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGly 702
 |||||
 Db 2073 TTTCTTGTGCACGACACTTAAATATGCTGATATCATCTTCCTTAAGAACTGAATTAATA 2132
 QY 702 sAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLe 722
 |||||
 Db 2133 GGGCATCAGAGATGTCGGGAGCGGTATCAATGATGCTTCCGTGAATGACAGACAGCT 2192


```

? PRIOR FILING DATE: 2000-05-03
? NUMBER OF SEQ ID NOS: 11
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 3
? LENGTH: 2415
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: This degenerate sequence encodes the amino acid
? OTHER INFORMATION: sequence of SEQ ID NO:2.
? NAME/KEY: misc_feature
? LOCATION: (1)...(2415)
? OTHER INFORMATION: n = A,T,C or G
? LS-09-978-385-3

```

Alignment Scores:

Pred. No.:	0	Length:	2415
Score:	3509.00	Matches:	644
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	161
Query Match:	81.78%	Indels:	0
DB:	9	Gaps:	0

US-09-978-385-2 (1-805) x US-09-978-385-3 (1-2415)

QY	1	MetSerSerSerSerThrLeuLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr	20
	11		
Db	1	ATGMSNMSNMSNWSMTGGTYTNTYTNTWMSNTNTNGTNGCNGTNAACGGCCARWSNAC	60
QY	21	IleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe	40
	11		
Db	61	ATHGARGARCARCGCNAARACWTYYTNTGAYAAFTTYYAAAYCAYGARGCGNARGAYTMTTY	120
QY	41	TyrGlnSerSerLeuAlaSerThrPAsnTyrAsnThrAsnIleThrGluGluAsnValGln	60
	11		
Db	121	TATCARMSNMSNTYNGCWSMTGGAAATYAAAYCAACNAATYTHACNGARGARAAAGTTCAR	180
QY	61	AsnMetAsnAsnAlaGluAspLysTrpSerAlaPheLeuIlysgGluGlnSerThrLeuAla	80
	11		
Db	181	AAATGAAATAYAGCGNGGAGAAATATGCGSNCCNTTYYTNAARGARCARWSNACWTYNGCN	240
QY	81	GlnMetYyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu	100
	11		
Db	241	CARATGATAYCCNTYNCARGARATHCARAAAYTNAACNTNARATYNCARNTNCARGCYTN	300
QY	101	GlnGlnAsnGlySerSerValLeuSerGluAspLysSerIlysarGluAsnThrIleLeu	120
	11		
Db	301	CARCARAAAYGGNMSNWSNNGTYTNTWMSNGARCAAAARMSNAARNGNTYNAAYCAACNATHYT	360
QY	121	AsnThrMetSerThrIleYyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu	140
	11		
Db	361	AAAYCAACNATGMSNACNATHATYAWSNACNGNARCTNTGYAACCCNGAYAACCCNARGAR	420
QY	141	CysLeuLeuLeuGluProGlyLeuAsnGluIleMetClnAsnSerLeuAspTyrAsnGlu	160
	11		
Db	421	TGYTNTNTNTNGARCCNGGAYTNAAGAAATATATGCCNMAAYSNYNTNGAYTAAAYGAR	480
QY	161	ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyGlnLeuArgProLeuYyr	180
	11		
Db	481	MGNTNTGGCGNTGGGARWSNTGTGMSNMGSNARGCTNGGNARCARFTTMMGCCNTYNTAT	540
QY	181	GluGluTyrValValLeuLysPheAsnGluMetAlaArgAlaAsnHisTyrGlnAspTyrGly	200
	11		
Db	541	GARGARATAYGTNGTYNTNNAARAAAYGARATGCMNGMGCNNAACATYAYGARAYTAYG	600
QY	201	AspTyrTrpAlaArgGluAspTyrGluValAsnGlyValAspGlyLysTyrAspTyrSerArgGly	220
	11		
Db	601	GATATAYTGGMGNGNGNATYTAGARCTNAAAGSGNTGATGCTNATYATAYTAYWSNMGG	660
QY	221	GlnLeuIleGluAspValGluHisThrPheGlnGluIleLysProLeuTyrGlnHisLeu	240
	11		
Db	661	CARYTNTTHARGARGAYGTNGARCAVACWTYTGARGARATTHARCCNTTNTATYARCAIYTN	720

QY	241	HisAlaTyrValAlaGlyAlaValLeuMetAspAlaTyrProSerTyrIleLeuProIleGly	260
Db	721	CAVGCGNTATYGTNNMGCCAAATYTAATGATGCTATACCCMSNTATATHTMSNCCNATHGCG	780
QY	261	CysLeuProAlaHisLeuLeuGlyAspMetIrrpGlyTyrGhrpThrAsnLeuTyrSer	280
Db	781	TGYTNNCCNCCNCAATYTYTNGGNGAATGAGGGGMMGNTTYTGACNAAATYNTAATYSN	840
QY	281	LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln	300
Db	841	YTMACNGNCCNTTYGGGCAAAACCCAAATATGAAGTMAACNCAATGCTGTNCAATCAR	900
QY	301	AlaTrpAspAlaGlnArgIlePheLysGluAlaGlyLysPheValSerAlaMetValAspGln	320
Db	901	GCNTGGAGYGCNCARMGATHTTAAAGARGCNGARARHTTYTGTMTMSNCTNGCNTYN	960
QY	321	ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln	340
Db	961	CCNAAATGACNCAKCGNTTYTGGGAAAWSNATGTATMACNCAATCCNAGNAAATGNCAR	1020
QY	341	LysAlaValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet	360
Db	1021	AARGNGNTGTCATCCNACNCGTGGGATYTTNGSAAARGNCAATYTMGNTHTYNTATG	1080
QY	361	CysThrLysValThrMetAspAspPheLeuThrAlaHisIleGluMetCylHisIleGln	380
Db	1081	TGYACNARGNACNATGAGATYTYTNACNGCCNCAATGATGGGNCATHTCAR	1140
QY	381	TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyValAsnGluGlyPhe	400
Db	1141	TATGATGATGGCNTATYGGGNCNCAKCCNTTYTYTMMGNAAAATGGNCCNAAATYARGCNTTY	1200
QY	401	HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuTyrSer	420
Db	1201	CATGARGCGNTGNSNGARATHTATGMSYTTMWSGCGNCAKCCNAAKARATYTNAAKSN	1260
QY	421	IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu	440
Db	1261	ATHGANTYNTMMSNCCNGATYTYTCARGARGAATGAARACNGARATHTAATYTYTNTYN	1320
QY	441	LysGlnAlaLeuThrIleValIleGlyThrLeuProPheThrTyrMetLeuGlnLysTrpArg	460
Db	1321	AARARCGNTTNACNATHGNSNCAKCYTNCCCTTATACATATATGTNGARARARGGCGN	1380
QY	461	TrpMetValPheLysGlyLuiLeuProLysAspGlnTrpMetLysLysTrpGluMet	480
Db	1381	TGATGTGNTTYTARGGNGARATHTCCNAAAGATCATGATGAARATRGNGGGARATG	1440
QY	481	LysArgGluIleValGlyValIleGluProValProHisAspGluThrTyrCysAspPro	500
Db	1441	AARNGNARARHTGNTGNSGNTNGNARCCNGNCCNCAAGAGARACATATGATGATCCN	1500
QY	501	AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu	520
Db	1501	GCMSNTNTNTTYCAYGTMNSNAAYGATYTAWSNTTYATHTMGNTATATACNMGNACVTN	1560
QY	521	TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis	540
Db	1561	TATACATTTTCATTYCARRGARGCNTYNTTYTCARCGNCCNAAACATGARGGCCNTYTTCAY	1620
QY	541	LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu	560
Db	1621	AARTGYATYATHTMSNAAYSNACNARGCGNGNCCARARNTTYTAATATGTYTMMGYTN	1680
QY	561	GlyLysSerIleProThrPheLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsn	580
Db	1681	GGNAAARSNARCCNCGACNTYNGCNTYNGARAAYGNTGNSGNCNAAATAATATGAAAY	1740
QY	581	ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys	600
Db	1741	GTMNGNCCNTYNTNAATYATYTYGARGCCNTYNTTYACNTGTYTNAAAGATCARAAATAR	1800
QY	601	AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal	620

```

||||| ||||||| ||||||| ||||||| |||||||
Db 1801 AAYMSNTTYGTGNGTGMWSNACNGAYTGWSNCCNTATGCGNAYCAWMSNATPAKGTN 1860
Qy 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet 640
Db 1861 MGNATHMSNTYNAAMWSNCGYNTNGNGNAYARAGCNTATGARTGGAAYGAAYGARATG 1920
Qy 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgIntYrPheLeuLysValLysAsn 660
Db 1921 TAYYNTTYMGWSNMSWSNGTNGCNTAYGCNATGMCNCRATYTYTYYTAAAGTNAARAAY 1980
Qy 661 GlnMetIleLeuPheGluGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680
Db 1981 CARATGATHTTNTTGGNGARAGARCAATGTMGNGTNGCNMAAYTYTNAACCMGNATHMSN 2040
Qy 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgTrpGluVal 700
Db 2041 TTYAAYTYTYYGTNACNGCNCNNAARAAYGTNMGNGAYATHTATHCNMGNAACNGARGTN 2100
Qy 701 GlnLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2101 GARARGCNATHTMGNATGWSNMGWSNMGNAHAAYGAYGCTTYMGNTYNAAGAYATAY 2160
Qy 721 SerLeuGluPheLeuGluIleGlnProThrLeuGluGlyProProAsnGlnProProValSer 740
Db 2161 WSNYNTNGATTTTNGNATHTCARCCNACNTNGNCCNCAATCARCCNCCNCTNMSN 2220
Qy 741 IleTrpLeuIleValPheGluValAlaMetGluValIleValValGluIleValIleLeu 760
Db 2221 AHTGTGNTATHTGTTTGTGNGTNGTNTATGGNGTNAHTGNGTNGTNAHTGNTAHTN 2280
Qy 761 IlePheThrGluIleArgAspArgGlyLysLysAsnLysAlaArgSerGluLysAsnPro 780
Db 2281 AHTHTTACGNAHTMGNAITMGNAIYGNARAARAARAAYARAGCMGWSNMGNGARARATCCN 2340
Qy 781 TyrAlaSerIleAspIleSerLysGluLysAsnAsnProGlyPheGlnAsnThrAspAsp 800
Db 2341 TAYGCMWSNATHTGAYATHTMSNARAGNGARAYAYCCNGNTYCARAAYACNGAYAY 2400
Qy 801 ValGlnThrSerPhe 805
Db 2401 GTNCARACNMSNTTY 2415

RESULT 20
US-09-978-385-7
; Sequence 7, Application US/09978385
; Patent No. US20020177211A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Petrie, Charles
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME
; FILE REFERENCE: 99-24CI
; CURRENT APPLICATION NUMBER: US/09/978, 385
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/133,952
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 60/151,181
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/563,516
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2415
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: This degenerate sequence encodes the amino acid
; NAME/KEY: misc_feature
; LOCATION: (1)...(2415)

```

```

; OTHER INFORMATION: n = A,T,C or G
US-09-978-385-7
Alignment Scores:
Pred. No.: 0
Score: 2904.00
Percent Similarity: 71.18%
Best Local Similarity: 65.47%
Query Match: 67.68%
DB: 9
Gaps: 0

US-09-978-385-2 (1-805) x US-09-978-385-7 (1-2415)
Qy 1 MetSerSerSerSerTrpLeuLeuLeuSerLeuValAlaValThrAlaGlnSerThr 20
Db 1 ATGWSNMSNMSNMSNTGGTYNTNTYNTNWSNTYNTGTCGNTNACNCCNCCNARMSNTYN 60
Qy 21 IleGluGluGlnAlaLysThrPheLeuAspLysPheAsnIleGluAlaGluAspLeuPhe 40
Db 61 ACNGARARARAAGCNARACNTTYTNAAYATTAAYAACARCARCGNARAGAYTNTMSN 120
Qy 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluLysAsnValGln 60
Db 121 TAYCARMSNMSNTYNTGCMNSNTGAAATAYAAACNAAATHTACNARARARAGCGNCAR 180
Qy 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
Db 181 AARATGSGNARCGNCGNCCNAAARTGWSNCGCTTYTAYGARARARARMSNARACNGCN 240
Qy 81 GlnMetYrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
Db 241 CARMSNTTYMSNTYNTACGARATHCARACNCCNATHTAARBNCGARFYTNCRNCGNTN 300
Qy 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerTrpArgLeuAsnThrIleLeu 120
Db 301 CARARMSNCGSNMSNMSNCGNTYNTMSNCGNATTAARAATARARARTNAAYACNATHYN 360
Qy 121 AsnThrMetSerThrIleYrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
Db 361 AAYACNATGWSNACNATHTAYWMSNACNGNARAGTNGTAYAACNNAARAAYCCNARCAR 420
Qy 141 CysLeuLeuLeuGluProGlyLysAsnGluIleMetAlaAsnSerLeuAspLysArgGlu 160
Db 421 TGYTYNTYNTYNTGACNCGNGNTYNTGAGARATHATGCGNACNMSNACNGTYTAYAAWSN 480
Qy 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuYr 180
Db 481 MGNATNTGGCNGTGGGARGGNTGGMGNGCNGARCTNGNAAARARCTTNGNCCNTNTAY 540
Qy 181 GlnGluTyrValAlaLeuLysAsnGluMetAlaArgAlaAsnIleTyrGluAspLysGly 200
Db 541 GARGARTAYGNGTNTYNTNAARAAYAGARATGCGNMGNCNNAAYATAAYAGAYTAGVN 600
Qy 201 AspTyrTrpArgGlyAspLysGluValAsnGlyValAspLysGlyAspLysSerArgGly 220
Db 601 GAYTAYTGGMGNGNGAYTAYGARGCNGARGGNGCNGAYGNTAYATAAYTAAYMGNAAV 660
Qy 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240
Db 661 CARITNATHGARAGAYGTNGARBMGNACNTTGGCNGARHTHAARCCNTNTNTYRGARCATY 720
Qy 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db 721 CAYGCNATAYGTMGNMGNMAARATYATAGAYACNATYACCMNSNTAYATHTMSNCCNACGN 780
Qy 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpHisAsnLeuTyrSer 280
Db 781 TGYTYNTGCGNCNCAYYNTYNTGNGAYATGTGGGNGNGNTTYTGACNAAYYNTNTAYCCN 840
Qy 281 LeuThrValProPheGluGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 841 YTNACNGTNCNTTYGCGNARARACNAAATHTGAYGTNACNGAYTNCNATGATGAAAYCAR 900

```

```

QY 301 AlAtRpAsPaLaGlaGnArGliePheLeuSglUaLglUuSpPheValSerValGlyLeu 320
    |||||..... |||||..... |||||.....
Db 901 GGNtGGAGtGCGNARGNAtHTTtTCAGANGCAGARARtTtTtTGtMNSNGtNGtYN 960
QY 321 ProAsnMetThrGlnGlyPheTPrGluAsnSerMetLeuThrAspProGlyAsnValGln 340
    |||||..... |||||..... |||||.....
Db 961 CCNcATATGACNcARGNtTtTGCGNAAYASNAATGtTtNCNcARcNGcNGAYGcNGM 1020
QY 341 LysAlaValAlcHisProThrAlaTTPAsPheGlyLysGlyAspPheArgIleuMet 360
    |||||..... |||||..... |||||.....
Db 1021 AARtTtNTtTGtCAcCNcACNcNGtGGAYtTGNGcAYtTGNGcAYtTtMNAARARtG 1080
QY 361 CysThrLysValThrMetAspPheLeuThrAlaHisGlnMetGlyHisIleGln 380
    |||||..... |||||..... |||||.....
Db 1081 TGtACNcARGNtMNAATGGAAYtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTt 1140
QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPhe 400
    |||||..... |||||..... |||||.....
Db 1141 TAYGAYATGCGtTtAYGcNMGNCARcCNTtTtTtTtTtTtTtTtTtTtTtTtTtTt 1200
QY 401 HisGluAlaValAlcGlyIleuMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
    |||||..... |||||..... |||||.....
Db 1201 CAAGAGcNGtNGcNARtHATGtMNSNGcNcACNcNcARcAYtTtNARtMNSN 1260
QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440
    |||||..... |||||..... |||||.....
Db 1261 ATtGAYtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTt 1320
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTrrParG 460
    |||||..... |||||..... |||||.....
Db 1321 AAcARcAGcNTtNcNcATtHcTNGcNcNcNTtTtTtTtTtTtTtTtTtTtTtTtTt 1380
QY 461 TrpMetValPheLysGlyIleuProLysAspGlnTrpMetLysTrrParGluMet 480
    |||||..... |||||..... |||||.....
Db 1381 TGtATGtGtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTt 1440
QY 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500
    |||||..... |||||..... |||||.....
Db 1441 AARtMNGARARtHcTNGcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNc 1500
QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520
    |||||..... |||||..... |||||.....
Db 1501 GCtMNSYtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTt 1560
QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHis 540
    |||||..... |||||..... |||||.....
Db 1561 TAYcATtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTt 1620
QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyLysLeuPheAsnMetLeuArgLeu 560
    |||||..... |||||..... |||||.....
Db 1621 AARtTGAYATtHMSNAAYASNAcNGcNcARARARtTtTtNARARtGtTtMNSYtN 1680
QY 561 GlyLysSerGluProThrPheLeuAlaGluAsnValAlaGlyAlaLysAsnMetAsn 580
    |||||..... |||||..... |||||.....
Db 1681 GGNtAYtMNSGARcCNTtGcNcARcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNc 1740
QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600
    |||||..... |||||..... |||||.....
Db 1741 GTtNARcCNTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTt 1800
QY 601 AsnSerPheValGlyTrrSerThrAspTrrSerProTyrAlaAspGlnSerIleLysVal 620
    |||||..... |||||..... |||||.....
Db 1801 AAYtMNSNTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTtTt 1860
QY 621 ArgIleSerLeuLysSerAlaLeuGlyLysPlyAlaTyrGluTrrPAsnAspAsnGlnMet 640
    |||||..... |||||..... |||||.....
Db 1861 MGNtATHMSYtTtNARtMNScNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNc 1920
QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660
    |||||..... |||||..... |||||.....
Db 1921 TtYtYtTtTtTtMNSMNSMNSMNSMNSMNSMNSMNSMNSMNSMNSMNSMNSMNS 1980
QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680

```

```

Db 1981 CARcNGtNcCNTtTtTtTtNcARcARGAYtGNMGtNcMNSGAYtTtNARcNcMNGtNcMNSN 2040
QY 681 PheAsnPheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700
    |||||..... |||||..... |||||.....
Db 2041 TtYtATtTtTtTtGtNcMNSNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNc 2100
QY 701 GluLysAlaIleArgMetSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
    |||||..... |||||..... |||||.....
Db 2101 GARGAYcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNc 2160
QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
    |||||..... |||||..... |||||.....
Db 2161 WSNYtNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNc 2220
QY 741 IleTrrPheIleValPheGlyValAlaMetGlyValIleValAlaGlyIleValIleLeu 760
    |||||..... |||||..... |||||.....
Db 2221 ATtHcGtYtNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNc 2280
QY 761 IlePheThrGlyIleArgAspArgLysLysAsnLysAlaArgSerGlyLysAsnPro 780
    |||||..... |||||..... |||||.....
Db 2281 ATtHcTtNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNc 2340
QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800
    |||||..... |||||..... |||||.....
Db 2341 TAYGAYtMNSNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNcNc 2400
QY 801 ValGlnThrSerPhe 805
    |||||..... |||||.....
Db 2401 GCNcNcARcNcMNSNTtTt 2415

```

Search completed: March 7, 2003, 11:03:34
Job time : 285 secs

